Guidance for Industry

Developing Medical Imaging Drugs and Biological Products

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication of the *Federal Register* notice announcing the availability of the draft guidance. Submit comments to Dockets Management Branch (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions on the content of the draft document contact (CDER) Robert K. Leedham Jr. 301-827-7510; or (CBER) George Q. Mills 301-827-5097.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
June 2000
Clin #

Guidance for Industry

Developing Medical Imaging Drugs and Biological Products

Additional copies of this Guidance are available from:

Office of Training and Communications
Division of Communications Management
Drug Information Branch, HFD-210
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane, Rockville, MD 20857
(Phone 301-827-4573)
Internet: http://www.fda.gov/cder/guidance/index.htm.

or

Office of Communication, Training and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
Internet: http://www.fda.gov/cber/guidelines.htm.
Fax: 1-888-CBERFAX or 301-827-3844
Mail: the Voice Information System at 800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
June 2000
Clin #

\\CDS018\REGAFF\!GUIDANC\3646dft2.DOC
06/19/00

Table of Contents

TABLE OF CONTENTS

Guidance for Industry[1]

Developing Medical Imaging Drugs and Biological Products

I. INTRODUCTION

This guidance is intended to assist developers of medical imaging drugs or biological products (*medical imaging agents*) in planning and coordinating their clinical investigations and preparing and submitting investigational new drug applications (INDs), new drug applications (NDAs), biologics license applications (BLAs), abbreviated NDAs (ANDAs), and supplements to NDAs or BLAs.

Medical imaging agents generally are governed by the same regulations as other drugs or biological products.[2] However, as described in this document, many medical imaging agents have special characteristics that can help guide developmental efforts. This guidance discusses some of these special characteristics and considers how development for medical imaging agents can be tailored to reflect those characteristics. Specifically, this guidance discusses the following issues:

1. Potential labeled indications for medical imaging agents and the nature of promotional materials for such claims[3]

2. Methods by which each of these labeled indications can be established

3. Special considerations in the clinical evaluation of efficacy

---

[1] This guidance has been prepared by the Division of Medical Imaging and Radiopharmaceutical Drug Products in the Center for Drug Evaluation and Research (CDER) and the Office of Therapeutics Research and Review in the Center for Biologics Evaluation and Research (CBER) at the Food and Drug Administration. This guidance represents the Agency=s current thinking on developing medical imaging drugs and biologics. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both.

[2] Sponsors developing medical imaging agents should be familiar with Agency regulations and guidances pertaining to the development of these products.

[3] The terms *claim*, *indication*, and *indication for use* are used interchangeably in this guidance.

4. Special considerations in the clinical evaluation of safety

A glossary of common terms used in diagnostic medical imaging is appended to the end of this document.

In response to the requirements of the Food and Drug Administration Modernization Act of 1997, FDA recently amended the drug and biologics regulations for one category of medical imaging agents by adding provisions for the evaluation and approval of in vivo radiopharmaceuticals used in the diagnosis or monitoring of diseases (64 FR 26657, May 17 1999). This guidance elaborates on the concepts contained in that final rule on radiopharmaceutical diagnostic products.

II. SCOPE — TYPES OF MEDICAL IMAGING AGENTS

This guidance applies to medical imaging agents that are used for diagnosis or monitoring and that are administered in vivo. Included are medical imaging agents used with medical imaging techniques such as radiography, computed tomography (CT), ultrasonography, magnetic resonance imaging (MRI), and radionuclide imaging. The guidance is not intended to apply to the development of in vitro diagnostic uses, or to therapeutic uses of these agents.

Medical imaging agents can be classified into two general categories:

A. Contrast Agents

Contrast agents improve the visualization of tissues, organs, and physiologic processes by increasing the relative difference of imaging signal intensities in adjacent parts of the body. Products include, but are not limited to (1) iodinated compounds used in radiography and CT; (2) paramagnetic metallic ions (such as ions of gadolinium, iron, and manganese) linked to a variety of molecules and used in MRI; and (3) microbubbles, microaerosomes, and related microparticles used in diagnostic ultrasonography.

B. Diagnostic Radiopharmaceuticals

As defined in 21 CFR 315.2 and 601.31 for diagnostic radiopharmaceuticals and as used in this guidance, a *diagnostic radiopharmaceutical* is (a) an article that is intended for use in the diagnosis or monitoring of a disease or a manifestation of a disease in humans and that exhibits spontaneous disintegration of unstable nuclei with the emission of nuclear particles or photons or (b) any nonradioactive reagent kit or nuclide generator that is intended to be used in the preparation of such an article. The FDA interprets this definition to include articles that exhibit spontaneous disintegration leading to the reconstruction of unstable nuclei and the subsequent emission of nuclear particles or photons (63 FR 28301 at 28303; May 22, 1998).

Diagnostic radiopharmaceuticals are radioactive drugs or biological products that contain a radionuclide that may be linked to a ligand or carrier.[4] These products are used in planar imaging, single photon emission computed tomography (SPECT), positron emission tomography (PET), or with other radiation detection probes.

Diagnostic radiopharmaceuticals used for imaging typically have two distinct components.

1. A radionuclide that can be detected in vivo (e.g., technetium-99m, iodine-123, indium-111). The radionuclide typically is a radioactive molecule with a relatively short physical half-life that emits radioactive decay photons having sufficient energy to penetrate the tissue mass of the patient. These photons can then be detected with imaging devices or other detectors.

2. A nonradioactive component that delivers the molecule to specific areas within the body. This nonradionuclidic portion of the diagnostic radiopharmaceutical often is an organic molecule such as a carbohydrate, lipid, nucleic acid, peptide, small protein, or antibody. In general, the purpose of the nonradioactive component is to direct the radionuclide to a specific body location or process.

III. INDICATIONS FOR MEDICAL IMAGING AGENTS

Medical imaging products are used clinically in many diverse ways, and this guidance outlines certain types of potential labeled indications for these agents. For example, some medical imaging agents are not intended to provide disease-specific information but are intended to characterize structural or functional manifestations common to several diseases. In such cases, the proposed indications for these products may refer to structural or functional assessments that are common to multiple diseases or conditions. In all cases, the effectiveness of a medical imaging agent is assessed by evaluating its ability to provide useful clinical information related to its proposed indications (see Section IV).

The labeled indications for medical imaging agents can fall within the following general categories:

- Structure delineation
- Functional, physiological, or biochemical assessment
- Disease or pathology detection or assessment
- Diagnostic or therapeutic patient management

These claims need not be mutually exclusive, and approval may be possible for claims other than those listed (see Section III.E). Each of these claims is described in the following sections as is the nature of

[4] In this guidance, the terms *ligand* and *carrier* refer to the entire nonradionuclidic portion of the diagnostic radiopharmaceutical.

promotional materials based on each of these labeled indications. Ways in which each of these labeled indications can be established are described in Section IV.

A. Structure Delineation

As described in the following sections, two types of labeled indications for structure delineation are possible: (1) locating and outlining normal anatomic structures and (2) distinguishing between normal and abnormal anatomy.

1. Locating and Outlining Normal Anatomic Structures

A medical imaging agent approved for this type of indication should be able to help locate and outline normal anatomic structures. The product also should help clarify the spatial relationship of the visualized normal structure(s) with respect to other body parts or structures.

Such a medical imaging agent can be developed to distinguish a normal structure that cannot be seen well with other imaging agents or modalities. For example, a contrast agent can be developed to image the normal parathyroid glands, which could help a surgeon plan and perform surgery for a mass in the thyroid gland. Products that help delineate normal anatomic variants also can be included here. An example of this type of product is an agent that delineates normal variants of coronary anatomy.

A medical imaging agent with this labeled indication enhances visualization of a normal anatomic structure or its variants and facilitates an understanding of the relationship of the normal visualized structure to other structures. Promotional materials based on this labeled indication should not imply that the product can be used to distinguish normal and abnormal anatomy, or that the product can be used to detect or assess disease or pathology. The materials should not imply that the product has been shown experimentally in adequate and well-controlled investigations to lead to more appropriate diagnostic or therapeutic management decisions in patients. These types of intended uses fall within other indications (see sections III.A.2, III.C, and III.D).

2. Distinguishing Between Normal and Abnormal Anatomy

A medical imaging agent approved for this type of indication should be able to help locate and outline both normal and abnormal anatomic structures. The agent also should help to clarify the spatial relationships of the normal and abnormal anatomic structure(s) with respect to other body parts or structures. This type of indication applies to situations where the mechanism by which the abnormal anatomy is visualized is sufficiently similar to the mechanism by which the normal anatomy is visualized. This type of indication does not apply to agents whose mechanism of visualization is dependent on the presence of an abnormality.

An example of this type of agent is one that nonspecifically enhances the airway lumen and that is being developed to help determine the distribution of bronchiectasis. The agent might be able to distinguish dilated bronchi from normal bronchi and categorize the bronchiectasis anatomically (e.g., as cylindric, sacculated, or fusiform). Similarly, a medical imaging agent that nonspecifically enhances the joint cavity might be developed to evaluate meniscal or ligamentous injuries of the knee. Products that help delineate anomalous variants of normal anatomy can also be included here (e.g., a product that helps define the anatomical relationships of a vascular sling that compresses the trachea or esophagus). In general, normal and abnormal structures are visualized by similar mechanisms in agents with this indication. If the mechanisms of visualization depend on the presence of the abnormality, a different indication would be more appropriate.

A medical imaging agent with this labeled indication helps distinguish between normal and abnormal anatomy or aids in the identification of variants or anomalies of normal anatomy. Promotional materials based on this labeled indication should not imply, beyond the description of the abnormal anatomy, that the product can be used to detect or assess disease or pathology, such as tumors or abscesses. Promotional materials should not imply that the product has been shown experimentally in adequate and well-controlled investigations to lead to more appropriate diagnostic or therapeutic management decisions in patients. These types of intended uses fall within other indications (see Sections III.C and III.D).

A medical imaging agent that is intended to delineate pathologic structures, such as tumors or abscesses, should seek a labeled indication of *disease or pathology detection or assessment* or *diagnostic or therapeutic patient management*, rather than this indication.

B. Functional, Physiological, or Biochemical Assessment

A medical imaging agent intended to provide functional, physiological, or biochemical assessment should be able to evaluate the function, physiology, or biochemistry of a tissue, organ system, or body region. Functional, physiological, and biochemical assessments are designed to determine if the value of a measured variable is normal or abnormal. This type of indication applies to agents used to detect either a reduction or magnification of a normal functional, physiological, or biochemical process. The indication of *functional, physiological, or biochemical assessment* is limited to assessment of functional, physiological, or biochemical processes when disturbances of these processes are common to several diseases or conditions and they are not diagnostic for any particular disease or condition. When these circumstances are not present, indications of *disease or pathology detection or assessment* or *diagnostic or therapeutic patient management* should be sought (see Sections III.C and III.D).

Examples of functional, physiological, or biochemical assessments include measurement of cardiac ejection fraction, assessment of regional cerebral blood flow, evaluation of myocardial wall motion, and assessment of anaerobic metabolites to evaluate tissue ischemia.

A claim of *functional, physiological, or biochemical assessment* should not be sought by sponsors who wish to develop a medical imaging agent for any of the intended uses listed here, because these types of uses fall within other indications (see Sections III.C and III.D):

! To establish a diagnosis by detecting or assessing the function, physiology, or biochemistry of a tissue, organ system, or body region

! To detect or assess an abnormality of function, physiology, or biochemistry that is diagnostic for a disease or condition

! To detect or assess an abnormality of function, physiology, or biochemistry that is diagnostic for a specific disease or condition in the defined clinical setting for which the test will be indicated and used

! To detect or assess functional, physiological, or biochemical processes that are not expressed by the normal organ system, tissue, or body part

A medical imaging agent with an indication of *functional, physiological, or biochemical assessment* facilitates assessments of function, physiology, or biochemistry. Promotional materials based on this labeled indication should not imply that the product can be used to detect or assess disease or pathology such as tumor or abscesses. The promotional materials should not imply that the product has been shown experimentally in adequate and well controlled investigations to lead to more appropriate diagnostic or therapeutic management decisions in patients. These types of intended uses fall within other indications (see Sections III.C and D).

For example, a medical imaging agent can be developed under the claim of *functional, physiological, or biochemical assessment* to assess cardiac ejection fraction, and it can be studied in subjects with a broad variety of representative cardiac diagnoses. Promotional materials for such a product based on this labeled indication can specify that the product facilitates the evaluation of ejection fraction, but these materials should not imply that the product can be used to establish a diagnosis or to determine the cause of myocardial dysfunction. However, if the medical imaging agent is being developed to detect or assess dose-related cardiac toxicity from anthracyclines (e.g., doxorubicin), or if promotional materials will be based on such a labeled indication, it should be studied under the indication *disease or pathology detection or assessment* in sufficient numbers of subjects who have received anthracyclines (see Section III.C).

C. Disease or Pathology Detection or Assessment

A medical imaging agent intended for disease or pathology detection or assessment should be able to assist in the detection, location, or characterization of a specific disease or pathological state in a defined clinical setting.[5] The medical imaging agent can be used alone or in combination with other diagnostic procedures to achieve this labeled indication.

Examples of medical imaging agents for which this type of indication would be appropriate include (1) an agent that can bind to a brain receptor and is being developed to detect or assess a specific neurological disease and (2) a radiolabeled monoclonal antibody that can attach to a tumor antigen and is being developed to detect or assess a tumor.

A medical imaging agent with this labeled indication facilitates detection or assessment of a specific disease or pathology in the defined clinical setting in which it was studied. Promotional materials based on this claim should not imply that use of the product has been shown experimentally in adequate and well-controlled investigations to lead to more appropriate diagnostic or therapeutic management decisions in patients or to improved clinical outcomes.[6] This type of intended use falls within another indication category (see Section III.D).

D. Diagnostic or Therapeutic Patient Management

A medical imaging agent that is intended for an indication of *diagnostic or therapeutic patient management* should improve diagnostic or therapeutic patient management decisions when used in a defined clinical setting.[7] To obtain this indication, adequate and well-controlled investigations should demonstrate experimentally that patient management decisions are, in fact, improved by use of the medical imaging agent (see Section IV.D.4). The medical imaging agent can be used alone or in combination with other diagnostic procedures to achieve this labeled indication.

Examples of medical imaging agents for which this type of indication would be appropriate include products that have been shown experimentally to improve decisions about whether patients should undergo diagnostic coronary angiography (i.e., use for diagnostic patient management) or be treated by tumor resection instead of with chemotherapy (i.e., use for therapeutic patient management). Labeling indications for these examples might include statements that the medical imaging agent is indicated *to help determine the need for coronary angiography or to assist in the evaluation of tumor resectability*.

---

[5] See Section IV.C for a definition of *defined clinical setting*.

[6] As used in this guidance, *clinical outcomes* refers to changes in patient symptoms, functioning, or survival.

[7] See Section IV.C for a definition of *defined clinical setting*.

Promotional materials based on this type of labeled indication may describe how the medical imaging agent assists in diagnostic or therapeutic patient management.

E. Multiple or Other Indications

The indication categories outlined above are flexible, and claims for medical imaging agents need not be mutually exclusive. A labeled indication can include several indication categories. For example, a diagnostic radiopharmaceutical could be developed as an aid in the diagnosis of lung cancer for a labeled indication of *disease or pathology detection or assessment*. This diagnostic radiopharmaceutical could also be evaluated in subpopulations of patients with lung cancer for its ability to provide information that leads directly to appropriate therapeutic management decisions (e.g., based on test results, determining what combination of surgery, radiotherapy, and chemotherapy is most appropriate).

Structural and functional aspects of diseases or conditions sometimes are evaluated together with imaging in clinical practice (e.g., as during ultrasonography). In such cases, clinical studies should evaluate the effect of the imaging agent on assessments of both structure and function. For example, if appropriate clinical studies are performed, an ultrasound contrast agent used to assess blood-vessel patency could receive an indication both for *structural delineation* and *functional assessment*. In this case, imaging studies might be designed so that structures of blood vessels and any obstructions are evaluated with two-dimensional ultrasonographic imaging. The functional hemodynamic consequences of these obstructions might be evaluated with Doppler interrogation of the same vessels.

For claims that do not fall within the indication categories identified above (e.g., providing prognostic information), the applicant or sponsor should consult FDA on the nature of the desired labeled indication and how to establish effectiveness for it.

IV. DEMONSTRATING EFFICACY FOR MEDICAL IMAGING AGENTS

To establish an indication for a medical imaging agent, a sponsor or applicant should characterize the agent's clinical usefulness and demonstrate that the information provided is valid and reliable.[8] Clinical studies should be performed in defined clinical settings that reflect the proposed indications. These overarching principles are discussed in this section, as are the methods of establishing effectiveness for specific indications.

---

[8] As used in this guidance, *validity* is a global concept that encompasses the quality of bias. Valid measurements are close to the *truth* (have small bias). *Reliability* is a concept that encompasses the quality of precision. Reliable measurements are reproducible (have small variance).

A. Clinical Usefulness

The effectiveness of a medical imaging agent is assessed by evaluating the agent's ability to provide useful clinical information related to its proposed indication.[9] A medical imaging agent that is *clinically useful* provides accurate and reliable information that adds to the appropriateness of diagnostic or therapeutic management, contributes to beneficial clinical outcome, or provides accurate prognostic information.

Depending on the specific indication, clinical usefulness can be established directly, indirectly, or historically. Clinical usefulness should be established directly for indications in which it cannot be established indirectly or historically. For example, clinical usefulness should be established directly for a medical imaging agent that has been shown in a research setting to bind specifically to particular receptors, but where it has not yet been established that evaluation of such binding adds to the appropriateness of diagnostic or therapeutic management, contributes to beneficial clinical outcome, or provides accurate prognostic information. Clinical usefulness can be established indirectly in some cases, such as when it is reasonable to infer that the test results lead to more appropriate management. For example, if a product is able to establish the diagnosis of early breast cancer, the clinical benefit of the use of this product can be inferred because treatment options are available for this stage of the disease (i.e., clinical usefulness has been established indirectly). Finally, clinical usefulness can be established historically when knowledge about the variable under study provides for an established clinical benefit. For example, medical imaging agents used to detect abdominal masses that need further evaluation, or medical imaging agents used to determine cardiac ejection fraction have clinical benefit that has been established historically in the medical literature. In such situations, clinical usefulness can be documented by a critical and thorough analysis of the medical literature and any historical precedents.

Test information that is inaccurate or unreliable can detract from appropriate management decisions, beneficial clinical outcomes, or accurate prognostic information. Therefore, assessments of clinical usefulness should weigh the possible benefits of the test information against its possible detrimental consequences. In such assessments of clinical usefulness, the possible benefits and possible detrimental consequences should be evaluated both for their quality and quantity.[10] For example, a moderate benefit from correct diagnoses in many patients could be offset by the significant detrimental consequences of incorrect diagnoses in a few patients (see Section X).

In some cases, information derived solely from a test with a medical imaging agent can be used to alter diagnostic or therapeutic management appropriately or to improve clinical outcome.

---

[9] 21 CFR 315.5(a) and 21 CFR 601.34(a)

[10] In decision analysis, the product of quality and quantity is termed *utility*.

However, in many cases medical imaging agents are used with other diagnostic information or other diagnostic tests. In all cases, tests with a medical imaging agent should contribute clinically useful information.

Accordingly, sponsors seeking indications of *structure delineation* or *functional, physiological, or biochemical assessment* should document how the medical imaging agent contributes information that is clinically useful. For an indication of *disease or pathology detection or assessment*, identification with sufficient validity and reliability of a disease or condition is adequate to demonstrate clinical usefulness provided that it is reasonable to infer that the test results lead to more appropriate management. For an indication of *diagnostic or therapeutic patient management,* experimental demonstration that the use of the medical imaging agent improves diagnostic or therapeutic decisions is sufficient to demonstrate clinical usefulness.

In addition, for a contrast agent to be considered clinically useful, the product used in combination with an imaging device should provide useful information or other advantages (such as improved imaging time or convenience) beyond that obtained by the imaging device alone. Imaging with the contrast agent should add value when compared to imaging without the contrast agent.

A plan for establishing clinical usefulness should be incorporated into the development plan of a medical imaging agent. In general, clinical usefulness should be evaluated prospectively in the principal clinical studies of efficacy (e.g., by incorporation into phase 3 protocols).

B. Validity of Information

The validity of information provided by a medical imaging agent generally should be established in adequate and well-controlled studies. In clinical studies, a medical imaging agent can be shown to provide valid information in at least two ways:

1. Comparing the results yielded by the medical imaging agent with the results of a truth standard (*gold standard*)[11]

2. Demonstrating that the use of the product contributes to beneficial patient outcomes

In instances where a truth standard does not exist or cannot be assessed practically, studies generally should be designed to evaluate the effects of the product on clinical outcomes. For example, clinical outcomes could be assessed in a study designed to evaluate the effects of the medical imaging agent on *diagnostic or therapeutic management* (see Section IV.D.4).

[11] See Glossary and Section VI.C.

Documenting structural or functional facts about humans or providing structural or functional data from appropriate animal studies also may be useful in establishing the validity of information provided by a medical imaging agent (see Sections IV.D.1 and IV.D.2).

C. Defined Clinical Settings

A *defined clinical setting* should reflect the circumstances and conditions under which the medical imaging agent is intended to be used. It delineates the patient population, relevant available medical and diagnostic data, and diagnostic questions that characterize the circumstances under which the medical imaging agent is intended to be used.[12] Generally, the choice of anticipated labeled indications will determine the clinical setting for the trials. In some cases, an appropriately designed trial may be able to include several clinical settings.

For example, a medical imaging agent that helps visualize duodenal ulcers could be developed for use in different defined clinical settings. The agent might be developed for one or more of the following *indications*: to detect duodenal ulcers in patients with gastrointestinal bleeding, to confirm the presence of suspected duodenal ulcers in patients with equivocal findings on radiographic examination of the upper gastrointestinal tract, to evaluate healing of duodenal ulcers in patients after initial treatment, or to help determine whether patients with duodenal ulcers should undergo surgery or remain on maintenance medical therapy. Similarly, the defined clinical setting of a screening evaluation differs from settings in which symptomatic individuals with physical findings are evaluated. For example, the setting in which otherwise asymptomatic healthy men undergo screening for prostatic cancer differs from a setting in which men with urinary symptoms and physical findings are evaluated for this condition.

The circumstances and conditions under which the medical imaging agent is intended to be used should be evaluated in clinical trials and can be described in the labeling using the following mechanisms.

1. Specifying aspects of the medical history and physical examination that are pertinent for determining the likelihood of the disease or condition that is in question. For example, a medical imaging agent intended to detect breast cancer might be evaluated for use in the assessment of (1) otherwise healthy women over 40 years of age, (2) women presenting with palpable breast masses, or (3) women with a family history of breast cancer.

[12] Note that use of a *defined clinical setting* in studies of medical imaging agents also tends to anchor both the *pretest probability* and the *spectrum* (e.g., severity or stage) of the disease or condition under study. Thus, when evaluated in a defined clinical setting, diagnostic performance measures that vary with the pretest probability of the disease or condition (e.g., positive and negative predictive values, accuracy), or that vary with the spectrum of the disease or condition (e.g., sensitivity, specificity, positive and negative predictive values, accuracy) tend to take on values that are relatively constant for that defined clinical setting.

2. Specifying a patient population that is at a particular step in the diagnostic sequence. For example, a diagnostic radiopharmaceutical may be intended to evaluate patients in an emergency room with equivocal clinical and laboratory findings of a myocardial infarction, or to evaluate the location and extent of a myocardial infarction in patients with definitive findings.

3. Specifying any other diagnostic assessments that are to be performed in the evaluation of this patient population. This delineation should include describing how the medical imaging agent should be used with respect to other diagnostic tests or evaluations, including (1) whether the medical imaging agent is intended to be used together with, or as a replacement for, other diagnostic tests or modalities and (2) how the use of the medical imaging agent is influenced by the results of other diagnostic evaluations.

   For example, in the evaluation of suspected pulmonary embolism, a medical imaging agent could be developed either as a replacement for ventilation-perfusion scanning or as an adjunct to ventilation-perfusion scanning. If the medical imaging agent is developed to be an adjunct to ventilation-perfusion scanning, its intended use will likely be influenced by the scan results. For instance, it may be intended for use in patients with scan results that are *intermediate* and not for patients with low-probability or high-probability scans.[13] Such a medical imaging agent should be studied in sufficient numbers of patients with intermediate scan results.

Clinical trials should prospectively evaluate relevant hypotheses about the demarcated patient population in the clinical setting in which the medical imaging drug or biologic is intended to be used.

[13] For the purpose of this example, *intermediate* scans are those with likelihood ratios for the presence of pulmonary embolism that are greater than low-probability scans, but less than that of high-probability scans.

D. Establishing Effectiveness for Specific Indications

The effectiveness of medical imaging agents for specific claims should be established through adequate and well-controlled clinical studies. The following sections describe how each of the types of indications summarized in Section III can be established.

1. Structure Delineation

To provide adequate estimates of the validity and reliability of the medical imaging agent over the full range of conditions for which it is intended to be used, medical imaging agents intended for this indication should be evaluated in studies with appropriate representation of sufficient numbers of subjects (a) with and without abnormalities, including the full spectra of abnormality and normality and (b) with other conditions, processes, or diseases that could affect the interpretation of the imaging results (e.g., inflammation, neoplasm, infection, trauma). *Appropriate representation* means that the studies generally should include subjects that adequately represent the spectra of normality and abnormality expected in the population in which the agent will be used. Methods by which indications for *structure delineation* can be established are described below.

For example, clinical trials of an agent intended to assess bronchiectasis should include adequate numbers of subjects over the full range of disease severity (e.g., from no disease to severe disease, or from early to late disease), subjects with local or diffuse disease, and subjects with related pulmonary disorders (e.g., chronic bronchitis, pneumonia, asthma, cystic fibrosis). Sponsors should justify the inclusion or exclusion of selected subpopulations during clinical development.

a. Locating and Outlining Normal Anatomic Structures

An indication of *delineating normal anatomic structures* can be established by demonstrating in clinical studies that the medical imaging agent can reliably locate and outline normal anatomic structures and reliably clarify the spatial relationship of these structures to other body parts.

In clinical studies, the validity of the delineation generally should be demonstrated by comparing the performance of the medical imaging agent with that of a reference product or procedure of known high validity (i.e., a truth standard). Ideally, the high validity of this reference product or procedure should be thoroughly and critically documented before initiating phase 3 studies.

In cases when valid reference products or procedures are unavailable or cannot be used feasibly, the validity of the information obtained with the medical imaging agent can be demonstrated with clinical studies that document how the product provides information that is consistent with known anatomic and structural facts about the tissue, organ, or body part in question. The sponsor should discuss these anatomic and structural facts with the Agency and carefully delineate and document them prior to initiation of phase 3 studies.

b. Distinguishing Between Normal and Abnormal Anatomy

An indication for *distinguishing between normal and abnormal anatomy* can be established by demonstrating in clinical studies that the medical imaging agent can reliably locate and outline both normal and abnormal variations of an anatomic structure, and that the product is able to clarify the spatial relationships of the normal and abnormal anatomic structures with respect to other body parts or structures.

Appropriate nonclinical studies in relevant animal models, if available, could provide additional information to support indications for *structure delineation*.

2. Functional, Physiological, or Biochemical Assessment

This type of indication can be established by demonstrating in clinical studies that the medical imaging agent can reliably measure a function, or a physiological or biochemical process. These measurements generally should be validated by comparing the performance of the medical imaging agent with that of a reference product or procedure of known high validity (i.e., a truth standard). Ideally, the high validity of this reference product or procedure should be documented thoroughly and critically before its use in clinical studies.

These studies should provide a quantitative or qualitative understanding of how the measurement varies in normal and abnormal subjects or tissues, including the variable's normal range, distribution, and confidence intervals in these subjects or tissues. When possible, the minimum detectable limits and reproducibility of the measurement should be assessed.

To provide adequate estimates of the validity and reliability of a medical imaging agent over the full range of conditions for which it is intended to be used, medical imaging agents intended for this indication should be evaluated in studies with appropriate representation of sufficient numbers of subjects (a) with and without abnormalities, including the full spectra of abnormality and normality and (b) with other conditions, processes, or diseases that could affect the interpretation of the imaging results (e.g., inflammation, neoplasm, infection, trauma). *Appropriate representation* means that the studies should generally include subjects that adequately represent the spectra of normality and abnormality expected in the population in which the agent will be used.

For example, clinical trials of a medical imaging agent intended to assess regional cerebral blood flow should include sufficient numbers of subjects who adequately represent the full range of functional, physiological, or biochemical dysfunction (e.g., from minimal or no perfusion to luxury perfusion) and subjects with inflammatory, neoplastic, infectious, or traumatic intracranial processes. Sponsors should justify the inclusion or exclusion of selected subpopulations during clinical development.

The agent's pharmacology in the setting of various functional, physiologic, or biochemical processes also should be documented from appropriate studies in relevant animal species, if available, to help establish the validity of the information obtained with the agent. These studies might include approaches such as induction of pharmacologic perturbations in the system to be evaluated (e.g., administration of a specific receptor antagonist that results in altered binding of the medical imaging agent); correlation with other accepted means of measuring particular variables (e.g., evaluation of the cardiac ejection fraction by comparison to results obtained with radionuclide ventriculography); and in vivo or in vitro analyses (e.g., tissue autoradiography). Documentation should be obtained in at least one appropriate and relevant animal species, if available, in which the particular function, physiology, or biochemistry is sufficiently similar to that of humans.

For example, full biochemical characterization of rodent brains by tissue autoradiography may be appropriate for a medical imaging agent being developed to evaluate particular receptors within the central nervous system. Such characterization could include in vitro receptor binding studies aimed at determining the binding affinity and specificity of the medical imaging agent. Such characterization also could include in vivo pharmacologic characterization of the distribution and density of the receptor in rodents using the medical imaging agent, including studies assessing effects of receptor agonists and antagonists on the binding or localization of the medical imaging agent in the brain.

3. Disease or Pathology Detection or Assessment

An indication of *disease or pathology detection or assessment* can be established by demonstrating in a defined clinical setting that the medical imaging agent is able to identify or characterize the disease or pathology with sufficient validity and reliability. In this context, the term *validity* refers to the overall diagnostic performance of the product as measured by factors such as sensitivity, specificity, positive and negative predictive values, accuracy, and likelihood ratios. Having *reliability* in this context means that the overall diagnostic performance of the product has precision. The phrase *sufficient validity and reliability* means validity and reliability that are good enough to indicate that the product could be useful in one or more defined clinical settings.

Data demonstrating validity and reliability should be obtained from patients in defined clinical settings reflecting the proposed indications. Patients can present for diagnostic evaluation of a specific disease or condition in various clinical settings. Even though these patients may be under evaluation for the same disease or condition, the likelihood of presence of the disease or the spectrum of the disease (e.g., severity or stage of disease) may vary in different clinical settings. The clinical usefulness and the diagnostic performance of the medical imaging agent may differ in each clinical setting. Pooling of efficacy data across defined clinical settings may be of limited value, and the medical imaging agent should be separately evaluated in sufficient numbers of patients in one or more settings. For example, pooling of efficacy data obtained with a medical imaging agent from patients being evaluated for early, localized malignancy (one clinical setting) with data from patients with advanced metastatic malignancy (another clinical setting) may be of limited value because the diagnostic performance of the agent may differ in these settings.

For similar reasons, if a medical imaging agent is being developed to diagnose a particular disease, efficacy trials generally should enroll subjects in whom the disease status is unknown, but in whom specific aspects of the clinical presentation have led to the desire for more diagnostic information. That is, the trials should include the intended population in the appropriate clinical setting. Data from subjects known definitely to have (or to not have) the disease of interest may be of limited value because these subjects are not the intended population for use of the medical imaging agent and therefore do not represent the defined clinical setting in which the medical imaging agent will be used.[14] Such enrollment may generate biased estimates of diagnostic performance because of *spectrum bias*.

Therefore, the medical imaging agent should be evaluated in representative settings in which its use is proposed. An indication *for disease or pathology detection or assessment* may specify the defined clinical setting and specify that the medical imaging agent should be used in conjunction with other tests.

4. Diagnostic or Therapeutic Patient Management

An indication of *diagnostic or therapeutic patient management* can be established in clinical studies by demonstrating experimentally that, in a defined clinical setting, the test is useful in guiding appropriate patient management. *Appropriate patient management* means that diagnostic or therapeutic management decisions are validated as being proper based on the correct diagnosis of the patient or based on clinical outcomes. The correct diagnosis can be documented by comparison with valid assessments of actual clinical status (e.g., a histological diagnosis of malignancy),

[14] These data are analogous to data obtained from a case-control study.

through patient follow-up, or by evaluation of clinical outcomes. For this indication, specific hypotheses of how the medical imaging agent might be useful in diagnostic or therapeutic management should be delineated in the protocol. These hypotheses should be tested prospectively in the clinical study and should be evaluated with endpoints that assess the appropriateness of patient management on clinical outcomes.

Medical imaging agents can be developed for indications of *disease or pathology detection or assessment*, or *diagnostic or therapeutic management*, or both. A clarification of the distinction between these indications is appropriate. The indication of *disease or pathology detection or assessment* can be obtained by demonstrating, in a defined clinical setting, sufficient validity and reliability of the medical imaging agent to imply clinical usefulness.

The indication of *diagnostic or therapeutic management* will likely be more difficult to establish, given the same defined clinical setting. Generally, it will require prospectively designed trials with the objective of evaluating a specific hypothesis of how the medical imaging agent might be useful in diagnostic or therapeutic patient management in a defined clinical setting. The trials might include randomization (whether or not to receive the medical imaging agent), with an endpoint measuring appropriateness of management (given the ultimate correct diagnosis) or clinical outcome. Alternatively, all patients can receive the study drug or biologic if it is possible to determine both what the management would have been had the medical imaging drug or biologic not been used and what the management would be because of information provided by the medical imaging agent. The trials should demonstrate that management based on findings using the medical imaging agent is superior to management without use of the medical imaging agent. A *patient management* indication can specify that the medical imaging agent is to be used in conjunction with other tests to influence a patient management decision.

A medical imaging agent intended to identify unrecognized disease in asymptomatic individuals (e.g., used in a screening setting) may obtain an indication of diagnostic or therapeutic patient management if it can be demonstrated experimentally that use of the test decreases irreversible morbidity or mortality. However, absent such an experimental demonstration, an indication of disease or pathology detection or assessment for a medical imaging agent could be supported by providing existing data that show that early detection and treatment of the disease decreases irreversible morbidity or mortality. For example, an indication of disease or pathology detection or assessment may be supported in such a circumstance by (a) clinical studies that demonstrate that the screening test is reproducible and has adequate sensitivity and specificity for the disease or condition of interest when it is applied to the population for whom the agent is intended to be used and (b) sufficient documentation that therapy would be more effective when the disease or condition is detected early by the medical imaging agent than when it is detected later by usual clinical methods (see Section IV.D.3).

V. GENERAL CONSIDERATIONS IN THE CLINICAL EVALUATION OF MEDICAL IMAGING AGENTS

Many considerations in the clinical development of drugs and biological products are discussed in various ICH and FDA guidance documents,[15] and the principles described in these documents also apply when developing medical imaging agents. General developmental principles include, but are not limited to, demonstration of safety and efficacy; procurement of adequate dose-response, pharmacodynamic, and pharmacokinetic data to support approval and special issues, such as consideration of drug metabolites, drug-drug interactions, and effects in special populations.

These guidances also discuss issues related to trial design, conduct, analysis, and how to report the data from individual clinical studies. The principles described in these documents apply just as well to individual clinical studies of medical imaging agents. Some obvious relevant topics include determining study objectives and study design; selecting subjects; evaluating dosage; selecting control groups, numbers of subjects, and response variables (i.e., endpoints or outcome measures); identifying methods to reduce bias (e.g., by randomization and blinding), and identifying important issues in statistical analysis.

However, the development of medical imaging agents for diagnostic purposes may also raise issues somewhat different from those raised during the development of therapeutic drugs or biological products. These issues deserve special attention. The following sections discuss some issues that are particularly relevant to the development of medical imaging agents. Considering them during the product development process should increase the efficiency of the clinical development of these products.

A. Phase 1 Studies

Phase 1 studies[16] can include, but are not limited to, assessments of the safety of single, increasing doses of a drug or biologic and evaluations of human pharmacokinetics. Depending on the drug or biologic and its potential toxicities, phase 1 studies may begin in healthy adult subjects or in patients. Screening for potential human toxicities can include serial evaluations of clinical laboratory tests (e.g., hematology, clinical chemistry, urinalysis), other laboratory tests (e.g., electrocardiograms), and adverse events (see Section X). Pharmacokinetic evaluations should address the absorption, distribution, metabolism, and excretion of all components of the formulation and any metabolites. Sponsors are encouraged to consult with the appropriate

[15] See ICH efficacy guidances available on the Internet at http://www.fda.gov/cder/guidance/index.htm, or http://www.fda.gov/cber/guidelines.htm.

[16] See also guidance for industry, *Content and Format of Investigational New Drug Applications (INDs) for Phase-1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-Derived Products* (November 1995).

FDA review division on pharmacokinetic issues. Evaluation of a medical imaging agent that targets a specific metabolic process or receptor should include assessments of its potential effects on directly related functions.

For diagnostic radiopharmaceuticals, organ and tissue distribution data over time should be collected to optimize subsequent imaging protocols and calculate radiation dosimetry (see Section X.C). Whenever possible, pharmacokinetic and pharmacodynamic evaluations should be made not only for the diagnostic radiopharmaceutical itself, but also for the radionuclide and for the carrier or ligand. The effects of large doses of the diagnostic radiopharmaceutical (including the carrier or ligand and other vial contents) should usually be assessed. This can be achieved, for example, by administering large doses of the medical imaging agent with low specific activity, by administering the contents of an entire vial of the medical imaging agent (assuming that this approximates a worst-case scenario in clinical practice), or both.

B. Phase 2 Studies

Goals of phase 2 studies of medical imaging agents should include refining the agent=s clinically useful dose range or dosage regimen (e.g., bolus administration or infusion), answering outstanding pharmacokinetic and pharmacodynamic questions, providing preliminary evidence of efficacy, expanding the safety database, optimizing the techniques and timing of image acquisition, developing methods and criteria by which images will be evaluated, and evaluating other critical concepts or questions about the medical imaging agent.

While refining the agent’s clinically useful dose range or dosage regimen, sponsors should explore the consequences of dose (or dosage regimen) adjustment on image acquisition and on the safety or effectiveness of the administered product. Additional exploration that should be considered during these studies include adjusting the following if relevant:

- character and amount of active and inactive ingredients
- amount of radioactivity
- amount of nonradioactive ligand or carrier
- specific activity
- radionuclide that is used

Methods used to determine the comparability, superiority, or inferiority of different doses or regimens should be discussed with the Agency. To the extent possible, the formulation that will be used for marketing should be used during phase 2 studies. When a different formulation is used, bioequivalence and other bridging studies may help document the relevance of data collected with the original formulation.

Phase 2 studies should be designed to define the appropriate patient populations and clinical settings for phase 3 studies. To gather preliminary evidence of efficacy, however, both subjects with known disease (or patients with known structural or functional abnormalities) and subjects known to be normal for these conditions can be included in clinical studies. Methods, endpoints, and items on the case report form (CRF) that will be used in critical phase 3 studies should be tested and refined.

C. Phase 3 Studies

The goals of phase 3 efficacy studies typically are to confirm the principal hypotheses developed in earlier studies, demonstrate the efficacy and continued safety of the medical imaging agent, and validate instructions for use and for imaging in the population for which the agent is intended. The design of phase 3 studies (e.g., dosage, imaging techniques and times, patient population, and endpoints) should be based on the findings in phase 2 studies (see Section VI.B). The formulation intended for marketing should be used, or else bridging studies should be performed.

When multiple efficacy studies are performed, the studies can be of different designs.[17] To increase the extent to which the results can be generalized, the studies should be independent of one another and should use different investigators, clinical centers, and readers that perform the blinded image evaluations (see Section VI.B).

VI. ADDITIONAL CONSIDERATIONS IN THE CLINICAL EVALUATION OF EFFICACY

The following sections describe special considerations for the evaluation of efficacy in clinical trials for medical imaging agents, and they complement items discussed in Section IV. Adequate and well-controlled studies should be designed to reduce possible biases by incorporating design elements that include, but are not limited to, appropriate selection of subjects, appropriate blinding procedures, choice of appropriate endpoints, and use of suitable truth standards and reference tests (if relevant). Sufficient detail should be provided in the protocol and study report to permit adequate characterization of the study population, imaging procedures, and other elements in the design, conduct, and analysis of the study.

A. Selecting Subjects

Subjects included in critical clinical efficacy studies should be representative of the population in which the medical imaging agent is intended to be used. The protocol and study reports should specify the method by which patients were selected for participation in the study (e.g., consecutive subjects enrolled, random selection) to facilitate assessments of potential selection bias. Other issues in appropriate subject selection for different indications are discussed in Section IV.D.

---

[17] See guidance for industry, *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products* (May 1998).

Subject selection for indications of (a) *structure delineation*, or (b) *functional, physiological, or biochemical assessment* can in some cases be based on representative diseases that involve similar alterations in structure, function, physiology, or biochemistry if it appears that the results can be extrapolated to other unstudied disease states based on a known common process. Appropriate models should be selected on a case-by-case basis. Data to justify inclusion of a particular disease should be thoroughly documented, as should the data to support why the results obtained from the models can be extrapolated to other diseases.

To aid subsequent clinical use of the medical imaging agent, the pretest odds and pretest probabilities of disease should be estimated for all subjects after enrollment, but before any trial results are made available. Whenever possible, these odds and probabilities should be derived from prespecified criteria of disease (e.g., history, physical findings, results of other diagnostic evaluations) according to prespecified algorithms. The estimated pretest odds and probabilities of disease should be compared with the pretest odds and probabilities actually observed in the studies.

B. Imaging Conditions and Image Evaluations

Medical imaging agents are used with many imaging modalities, and imaging data can be acquired, reconstructed, processed, stored, and displayed in numerous ways. Because of this heterogeneity, the sponsor may want to customize general recommendations delineated below for imaging and image evaluation in clinical trials to fit a specific medical imaging drug, biologic, or imaging modality.

The following sections use the term *images* in a general way. For example, an *image* of the heart obtained with a diagnostic radiopharmaceutical or an ultrasound contrast agent may in some cases refer to a *set* of images acquired from different views of the heart (e.g., short-axis and long-axis views). Similarly, an *image* obtained with an MRI contrast agent may in some cases refer to a *set* of images acquired with different pulse sequences and interpulse delay times.

1. Imaging Conditions

Conditions for using a medical imaging agent with its corresponding imaging device should be evaluated during early product development. Subsequently, the imaging conditions that are anticipated for clinical use should be employed in the principal efficacy trials. For example, the effects of changes in relevant imaging conditions (e.g., timing of imaging after product administration, views, instrument settings, patient positioning) on image quality and reproducibility, including any limitations imposed by changes in such conditions, should be evaluated in early product development. Subsequent principal efficacy trials should substantiate and may refine these conditions for use. Appropriate imaging conditions, including limitations, can be described in the product labeling.

2. Methods and Criteria for Image Evaluation

Methods and criteria for image evaluation (including criteria for image interpretation) should be evaluated in early product development. Subsequently, the methods and criteria that are anticipated for clinical use should be employed and substantiated in the principal efficacy trials. For example, early clinical trials might compare ways in which regions of interest on images are selected or ways in which an organ will be subdivided on images for purposes of analysis. The most appropriate of these methods could then be incorporated into the protocols of the principal efficacy trials. Similarly, early clinical trials might evaluate which objective image features (e.g., lesion conspicuity, relative count rate density) appear to be most affected by the medical imaging agent and which of these are most useful in image interpretation, such as making a determination of whether a mass is benign or malignant (see Section VI.B.3). The most appropriate of these criteria for image evaluation could then be incorporated into the protocols of the principal efficacy trials. Appropriate methods and criteria for image evaluation, including limitations, can be described in the product labeling.

Sponsors should seek FDA comment on the designs and analysis plans for the principal efficacy trials before they are finalized (see Section V.C). In addition, the following elements should be completed and submitted to the IND before the principal efficacy studies enroll subjects:

- Proposed indications for use
- Protocols for the principal efficacy trials
- Investigators' brochure
- CRFs to be used by on-site investigators
- Plan for blinded image evaluations[18]
- CRFs to be used by the blinded readers
- Statistical analysis plan
- Plan for on-site image evaluation and intended use of such evaluation in patient management, if any

Sponsors should submit a single comprehensive statistical analysis plan for each principal efficacy study. This statistical analysis plan should be part of the study protocol, should include the plan for blinded image evaluations, and should be submitted to the protocol before images have been collected.

3. Steps in Image Evaluation

[18] *Blinded* image evaluations may also be referred to as *masked* or as *uninformed* image evaluations.

Broadly speaking, the evaluation of medical images consists of two distinct steps: assessing objective image features and interpreting findings on the image.

a. Assessing Objective Image Features

Objective image features are attributes on the image that are either visually perceptible or that can be detected with instrumentation. Examples of objective image features include signal-to-noise ratios; degree of delineation; extent of opacification; and the size, number, or density of lesions. Objective image features can be captured on scales that are continuous (e.g., the diameter of a mass), ordinal (e.g., a feature can be classified as definitely increased, probably increased, neither increased nor decreased, probably decreased, definitely decreased), or dichotomous (e.g., a feature can be classified as present or absent). Thus, a reader who classifies the intensity of radionuclide localization in a tissue as decreased, similar, or increased compared to the surrounding tissue is describing an objective image feature on a qualitative (ordinal) scale. That is, a feature can be an objective one, even if the scale being used to measure it is qualitative.

Medical imaging agents have their intended effects by altering objective image features. Both the nature and location of such changes on the image should be documented fully during image evaluations in clinical trials intended to demonstrate or support efficacy. Such documentation should include information not only on changes that are intended and desirable, but also on changes that are unintended or undesirable. For example, a diagnostic radiopharmaceutical intended for cardiac imaging also might localize in the liver, thereby obscuring visualization of parts of the heart. Such effects should be documented.

When possible, it is often desirable to perform both a qualitative visual evaluation of images as well as a quantitative analysis of images with instrumentation. For example, a quantitative image analysis with instrumentation could help corroborate visual findings, and such an analysis could provide important evidence that supports the efficacy of the medical imaging agent. However, a quantitative image analysis with instrumentation by itself may not be sufficient to establish efficacy of the medical imaging agent, such as in cases where images are not intended (or not likely) to be evaluated quantitatively with instrumentation in clinical practice. In such cases, studies should establish that visual image evaluations are capable of discerning changes caused by the medical imaging agent on the pertinent objective image features.

b. Image Interpretation

An *image interpretation* is the explanation or meaning that is attributed to objective image features. Interpretations of image features should be supported by objective quantitative or qualitative information derived from the images. For example, the interpretation that cardiac tissue seen on an image is infarcted, ischemic, or normal might be supported by objective image features such as the extent and distribution of localization of the medical imaging agent in the heart (e.g., increased, normal, decreased, or absent), the time course of such localization, and how these features are affected by exercise or pharmacologic stress.

4. Endpoints in Trials

Medical imaging agents can be developed for many different reasons, such as to help make a diagnosis, to alter patient management, to ascertain the severity of a condition, or to determine the prognosis of an illness. In clinical trials designed to establish the efficacy of a medical imaging agent, a principal objective generally should be to evaluate the effects of imaging with the agent on one or more of such clinically meaningful items. Accordingly, as summarized below, the primary endpoints (response variables) in clinical trials designed to establish or support the efficacy of a medical imaging agent usually should be directly related to such clinically meaningful items.

a. Image Interpretations as Endpoints

Image interpretations often have clinical implications, and such interpretations can be incorporated into the primary endpoint in clinical trials designed to establish or support the efficacy of a medical imaging agent. For example, the primary endpoint (response variable) of a trial for a medical imaging agent intended to aid in the diagnosis of lung cancer, such as for a claim of *disease or pathology detection or assessment*, might be the proportion of subjects with and without the disease who are properly classified. In this example, the interpretation that a pulmonary lesion seen on an image is benign or malignant has direct clinical meaning and can be incorporated into the primary endpoint.

b. Objective Image Features as Endpoints

When the clinical implications of particular objective image features are apparent, the objective imaging features can be incorporated into the primary endpoint. For example, in a study of a medical imaging agent intended for brain imaging, the ability to identify the presence or absence of cranial masses on images has direct clinical meaning and might be incorporated into the primary endpoint to serve as the primary basis for the indication for the product (e.g., the medical imaging agent is indicated for detecting cranial masses in patients in a particular defined clinical setting).

However, in some cases the clinical implications of particular objective image features may not be readily apparent without additional interpretation. In these cases, the objective image features generally should serve as secondary imaging endpoints. For example, the finding that a medical imaging agent alters the conspicuity of masses differentially could lead to the interpretation that specific masses are benign or malignant; acute or chronic; inflammatory, neoplastic, or hemorrhagic; or lead to some other clinically meaningful deductions. Such interpretations can be incorporated into the primary endpoint and can serve as the primary basis for the indication for the product. In such a case, however, the objective image feature of lesion conspicuity might be designated more appropriately as a secondary imaging endpoint.

c. Subjective Image Assessments as Endpoints

Subjective image assessments, if included as endpoints in clinical trials intended to demonstrate or support efficacy of a medical imaging agent, should be linked to objective image features so that the objective basis for such assessments can be understood. *Subjective image assessments* are assessments that are perceptible only to the reader. Such assessments are not visually perceptible and cannot be detected with instrumentation. For example, a conclusion that use of a medical imaging agent alters *diagnostic confidence* is a subjective assessment as is the conclusion that a medical imaging agent provides *more diagnostic information*. Subjective image assessments can be difficult to validate and replicate, and the possibility that substantial bias has been introduced into such assessments often cannot be excluded. Accordingly, subjective image assessments generally should not be used as primary imaging endpoints.

d. Clinical Outcomes as Endpoints

Clinical outcomes, such as measurement of symptoms, functioning, or survival, are among the most direct ways to measure clinical benefit. Accordingly, clinical outcomes can serve as primary endpoints in trials of medical imaging agents. For example, the primary endpoint of a trial of a medical imaging agent intended for a claim of therapeutic patient management in patients with colon cancer might be a response variable that measures changes in symptoms, functioning, or survival (see Section IV.D.4).

5. Case Report Forms

Case report forms (CRFs) in trials of medical imaging agents should prospectively define the types of observations and evaluations for investigators to record. In addition to data that are usually recorded in CRFs (e.g., inclusion/exclusion criteria, safety findings, efficacy findings), the onsite investigator's CRF for a medical imaging agent should capture the following information:

- The technical performance of the diagnostic radiopharmaceutical used in the study, if any (e.g., specific activity, percent bound, percent free, percent active, percent inactive)

- The technical characteristics and technical performance of the imaging equipment (e.g., background flood, quality control analysis of the imaging device, pulse height analyzer)

- Methods of image acquisition, output processing, display, reconstruction, and archiving of the imaging study

6. CRFs for Image Evaluation

Imaging CRFs should be designed to capture imaging endpoints, including objective features of the images as well as the location and interpretation of any findings. Interpretations of image features should be supported by objective quantitative or qualitative information derived from the images. Image interpretations should be recorded as distinct items from the assessments of the objective image features. Items on the CRFs for image evaluation should be carefully constructed to gather information without introducing a bias that indicates the answer that is being sought. The proposed labeled indication should be clearly derived from specific items in the CRF and from endpoints and hypotheses that have been prospectively stated in the protocol.

7. Blinded Imaging Evaluations

Image evaluations should be designed to demonstrate that the medical imaging agent provides useful clinical information about its proposed indications for use (see Section IV.A). Moreover, image evaluations should be designed to demonstrate that the specific effects of the medical imaging agent, as manifested in the images, provide such information reproducibly and apart from other possible confounding influences or biases. Thus, as described and defined below, blinded image evaluations by multiple independent readers should be performed in the principal efficacy studies of diagnostic radiopharmaceuticals or contrast agents.

Specifically, a *fully blinded image evaluation* or an *image evaluation blinded to outcome* by independent readers generally should serve as the principal image evaluation for demonstration of efficacy to support licensing of medical imaging agents.[19] Such image evaluations can be performed through *sequential unblinding*. Both primary and secondary imaging endpoints should be evaluated with such image evaluations whenever they are to be used to demonstrate or support efficacy. For image evaluations intended to demonstrate efficacy, the nature and type of information available to the readers should be discussed with FDA before the trials are initiated.

In addition to the items outlined in the sections below, plans for blinded image evaluations should include the following elements:

[19] See Section VI.B.8 for a definition of *independent readers*.

- The protocol should clearly specify the elements to which readers are blinded.

- Meanings of all endpoints should be clearly understood for consistency. Terms to be used in image evaluation and classification should be defined explicitly in the image evaluation plan, including such terms as *technically inadequate*, *uninterpretable*, *indeterminate*, or *intermediate.* Blinded readers can be trained in scoring procedures using sample images from phase 1 and phase 2 studies.

- Images should be masked for all patient identifiers.

- Blinded readers generally should evaluate images in a random sequence. *Randomization* of images refers to merging the images obtained in the study (to the fullest degree that is practical) and then presenting images in this merged set to the readers in a random sequence. For example, when images of several diagnostic radiopharmaceuticals read by the same criteria are being compared to establish relative efficacy (e.g., a comparison of a test drug or biologic to an established drug or biologic), the readers generally should evaluate individual images from the merged set of images in a random sequence.

a. Fully Blinded Image Evaluation

At a minimum during a *fully blinded image evaluation*, readers should not have any knowledge of the following types of information:

- Results of evaluation with the truth standard, of the final diagnosis, or of patient outcome

- Any patient-specific information (e.g., history, physical exam, laboratory results, results of other imaging studies)

In some cases, general inclusion and exclusion criteria for patient enrollment, other details of the protocol, or anatomic orientation to the images also should not be provided to the readers.

During a *fully blinded image evaluation* in studies where images obtained by different treatments are being evaluated, readers should not have knowledge of treatment identity, to the greatest extent to which that is possible.[20] For example, in a comparative

[20] This is the common meaning of *blinding* in therapeutic clinical trials. See *E8 General Considerations for Clinical Trials* (ICH) (December 17, 1997), and *E9 Statistical Principles for Clinical Trials* (ICH) (September 16, 1998).

study of two or more medical imaging agents (or of two or more doses or regimens of a particular medical imaging agent), the blinded readers should not know which agent (or which dose or regimen) was used to obtain a given image. For contrast agents, this also can include lack of knowledge about which images were obtained before product administration and which were obtained after product administration, although sometimes this is apparent upon viewing the images. In cases where the instructions for image evaluation differ according to treatment (e.g., as might be the case when images are obtained using different imaging modalities), blinding the readers to treatment identity may be infeasible.

b. Image Evaluation Blinded to Outcome

As in a *fully blinded image evaluation*, readers performing an *image evaluation blinded to outcome* should not have any knowledge of the results of evaluation with the truth standard, of the final diagnosis, or of patient outcome.

However, in an *image evaluation blinded to outcome* the readers may have knowledge of particular elements of patient-specific information (e.g., history, physical exam, laboratory results, or results of other imaging studies). In some cases, the readers also may be aware of general inclusion and exclusion criteria for patient enrollment, other details of the protocol, or anatomic orientation to the images. The particular elements of which the reader will have information should be standardized for all patients and defined prospectively in the clinical trial protocol, statistical plan, and the blinded image evaluation plan.

In studies where images obtained by different treatments are being evaluated (including *no treatment*, such as in unenhanced image evaluation of a contrast agent), the readers should not have knowledge of treatment identity, to the greatest extent to which that is possible (see Section VI.B.7.a).

c. Sequential Unblinding

In *sequential unblinding*, readers typically evaluate images with progressively more information (e.g., clinical information) on each read. Sequential unblinding might be used to provide incremental information under a variety of conditions that may occur in routine clinical practice (e.g., when no clinical information is available, when limited clinical information is available, and when a substantial amount of information is available). This can be used to determine when or how the test agent should be used in a diagnostic algorithm. A typical s*equential unblinding* image evaluation is a three-step process.

- A fully blinded image evaluation is performed. This evaluation is recorded and locked in a dataset by methods that can be validated. In a *locked* dataset, it should not be possible to alter the evaluation later when additional information is available, or if input is received from the clinical investigators, other readers, or the sponsor.

- An image evaluation blinded to outcome is performed. This evaluation is recorded and locked in the dataset.

- To determine diagnostic performance of the imaging agent, the result of the above two blinded evaluations are compared to the results of evaluation with the truth standard (or of the final diagnosis, or of patient outcome).

Such sequential unblinding can be expanded to include other types of image evaluations where additional clinical information is provided to the readers. If sequential unblinding is used, the protocol should specify the hypothesis that is to be evaluated at each step. Also, the protocol should specify which image evaluation will be the primary one for determining efficacy.

d. Unblinded Image Evaluations

In an *unblinded image evaluation,* readers are aware of the results of patient evaluation with the truth standard, of the final diagnosis, or of patient outcome. Unblinded readers also typically are aware of patient-specific information (e.g., history, physical exam, laboratory results, results of other imaging studies), of treatment identity where images obtained by different treatments (including no treatment) are being evaluated, of inclusion and exclusion criteria for patient enrollment, other details of the protocol, and of anatomic orientation to the images.

In trials intended to demonstrate or support efficacy, unblinded image evaluations can be used to show consistency with the results of fully blinded image evaluations or image evaluations blinded to outcome. However, unblinded image evaluations should not be used as the principal image evaluation for demonstration of efficacy. For example, unblinded readers may have additional information about patients that was not predefined in the clinical trial protocol. Such additional information may alter the readers' diagnostic assessments and may confound or bias the image evaluation by these readers. Blinded and unblinded image evaluations should use the same endpoints so that the results can be compared.

8. Independent Image Evaluations

As stated above, image evaluations should be performed by multiple independent blinded readers in trials intended to demonstrate efficacy of the medical imaging agent. Two events are independent if knowing the outcome of one event says nothing about the outcome of the other. Therefore, *independent readers* are readers that are completely unaware of findings of other readers (including findings of other blinded readers and onsite investigators) and are readers who are not otherwise influenced by the findings of other readers. To ensure that blinded reader's evaluations remain independent, each blinded reader's evaluation should be locked in the dataset shortly after it is obtained and before additional types of image evaluations are performed (see Section VI.B.7.c.i).

a. Consensus Image Evaluations

Consensus image evaluations (*consensus reads*) are image evaluations during which readers convene to evaluate images together. Consensus image evaluations can be performed after the individual readings are completed and locked. However, readers are not independent during consensus reads and therefore such reads should not serve as the primary image evaluation used to demonstrate efficacy of medical imaging agents. Although a consensus read is performed by several readers, it is actually a single image evaluation and does not fulfill the need for image evaluations by multiple blinded readers. As with the individual blinded evaluations, the consensus reads should be locked once obtained and before additional types of blinded readings are performed.

b. Repeated Image Evaluations by the Same Reader

In studies where readers evaluate the same image multiple times (e.g., as in sequential unblinding, or in readings designed to assess *intra*reader variability), the readings should be performed independently of one another to the fullest extent practical. This means that the readers should be unaware, to the fullest extent practical, of their own previous image findings and should not be otherwise influenced by their own previous findings.

Stated differently, such blinded reading sessions generally should be designed to decrease *recall bias*. For example, if an image evaluation blinded to outcome is performed after a fully blinded image evaluation as during sequential unblinding (see Section VI.B.7.c), different pages in the CRF should be used for the two types of image evaluation, and each image evaluation usually should be performed with sufficient time between readings to decrease recall and without reference to prior results.

9. Offsite and Onsite Image Evaluations

Offsite image evaluations are image evaluations performed at sites that have not otherwise been involved in the conduct of the study, and by readers who have not had contact with patients, investigators, or other individuals involved in the study. Trials intended to demonstrate or support efficacy generally should include offsite image evaluations that are performed at a limited number of sites (or preferably at a centralized site). In such offsite evaluations, it is usually easier to control factors that can compromise the integrity of the blinding image evaluations and to ensure that the blinded readers perform their image evaluations independently of other image evaluations. For example, offsite readers generally are not likely to have had any involvement with the patients, investigators, or other individuals involved in the study. They are therefore unlikely to become inadvertently unblinded to clinical data or to become unintentionally aware of the results of image evaluations by others, either of which could affect their image evaluations.

Onsite image evaluations are image evaluations performed by investigators involved in the conduct of the protocol or in the care of the patient. The term also can refer to blinded image evaluations performed at sites involved with the conduct of the study. Onsite investigators may have additional information about the patients that was not predefined in the clinical trial protocol. Such additional information may alter the investigators' diagnostic assessments and may confound or bias the image evaluation by the investigators. Therefore, onsite image evaluations usually should not be used as the principal image evaluation for demonstration of efficacy but generally should be regarded as supportive of the blinded image evaluations.

However, if onsite investigators blinded to *truth* (e.g., blinded to any test result that make up the truth standard, to the final diagnosis, and to patient final outcome as in an image evaluation blinded to outcome; see Section VI.B.7.b) are to perform image evaluations that are to be part of the demonstration of efficacy, then all clinical information available to the investigator at the time of the image evaluation should be clearly specified and fully documented. A critical assessment of how such information might have influenced the readings should be performed. In addition, an independent blinded evaluation that is supportive of the finding of efficacy should be performed.

10. Assessment of Interreader and Intrareader Variability

At least two blinded readers (and preferably three or more) should evaluate images for each study that is intended to demonstrate efficacy. This allows for an evaluation of the reproducibility of the readings (i.e., interreader variability) and provides a better basis for subsequent generalization of any findings. Ideally, each reader should view all of the images intended to demonstrate efficacy so that interreader agreement can be measured. In large studies, where it may be impractical to have every image read by each reader, a properly chosen subset of images can be selected for such duplicate image evaluations. Consistency among readers should be measured quantitatively (e.g., with the kappa statistic).

*Intra*reader variability should be assessed during the development of medical imaging agents. This can be accomplished by having individual blinded readers perform repeated image evaluations on some or all images (see Section VI.B.8.b).

11. Protocol and Nonprotocol Images

Images obtained in a clinical trial of a medical imaging agent can be classified either as protocol or nonprotocol images.

a. Protocol Images

For the purpose of this guidance, *protocol images* are images obtained under protocol-specified conditions and at protocol-specified time points with the goal of demonstrating or supporting efficacy. Efficacy evaluations should be based primarily upon the evaluations of such protocol images. Ideally, all protocol images (e.g., not just those images determined to be evaluable) should be evaluated by the blinded readers, including images of test patients, control patients, and normal subjects. Evaluation of the protocol images should be completed before other images, such as nonprotocol images, are reviewed by the readers (see Section VI.B.11.b).

In some cases where large numbers of images are obtained or where image tapes are obtained (e.g., cardiac echocardiography), sponsors have used image selection procedures. This generally is discouraged because the selection of images can introduce the bias of the selector. In cases where preselection is thought to be needed, the sponsor is encouraged to clearly identify and discuss the selection procedures with the appropriate Agency division before their implementation.

Sponsors should specify prospectively in protocols of efficacy studies how missing images (and images that are technically inadequate, uninterpretable or show results that are indeterminate or intermediate) will be handled in the data analysis. For example, images may be missing from analysis for many reasons, including patient withdrawal from the study, technical problems with imaging, protocol violations, and image selection procedures. Sponsors are encouraged to incorporate analyses in the statistical analysis plan that incorporate the principle of intention-to-treat, but which are adapted to a diagnostic setting (e.g., *intention-to-image* or *intention-to diagnose*).[21]

b. Nonprotocol Images

For the purpose of this guidance *nonprotocol image* refers to an image that is not a protocol image, as defined above (see Section VI.B.11.a). Nonprotocol images include those that have been generated under conditions or at time points that were not specified in the protocol. If such additional nonprotocol images are presented to the blinded readers, they should be presented to the readers only after they have made and locked their final reading of the protocol images.

[21] The *intention-to-treat principle* is defined as the principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a subject (i.e., the planned treatment regimen) rather than the actual treatment given. It has the consequence that subjects allocated to a treatment group should be followed up, assessed, and analyzed as members of that group irrespective of their compliance with the planned course of treatment (See ICH *E9*; P. 49597).

12. *Separate or Combined Image Evaluations*

As described in the following sections, separate image evaluations by independent, blinded readers should be performed in some studies intended to demonstrate or support efficacy of a medical imaging agent. Combined image evaluations by independent, blinded readers may also be useful in evaluating the efficacy of a medical imaging agent. Regardless of whether a separate or combined image evaluation is performed (or if both are performed), image evaluations usually should consist of blinded, randomized, independent readings that are designed to evaluate whether the medical imaging agent contributes additional useful information.

Performance of a separate image evaluation does not preclude performance of a combined image evaluation, and vice versa. Both types of image evaluations can be performed if desired. If multiple image evaluations are performed, however, the protocol should specify which image evaluation will serve as the primary evaluation and which image evaluations are secondary.

a. Separate Image Evaluations

In a *separate* image evaluation, a reader evaluates test images obtained from a patient independently of other test images obtained from that patient, to the fullest degree practical.[22] In such an evaluation, the reader generally should not be influenced by evaluations of other test images obtained from that patient, including any previous evaluations of test images performed by the same reader for that patient (see Section VI.B.8). In this context, other test images include those obtained under different conditions (e.g., with different medical imaging agents) or at different times with respect to agent administration. In other words, in a separate image evaluation, a reader evaluates each test image for a patient on its own merits without reference to, or recall of, any other test images obtained from that patient, to the fullest degree practical.

A separate image evaluation often can be performed by combining test images obtained under different conditions (or at different times) into an intermixed set. Images in this intermixed set can then be evaluated individually in random order so that multiple images are not viewed simultaneously, and so that images are not evaluated sequentially within patients. Alternatively, test images obtained under one condition (or at a particular time) can be evaluated individually in a random order, followed by an evaluation in random order of the individual test images obtained under different conditions (or at different times).

[22] In the special case where only two test images are being evaluated, a *separate* image evaluation may also be referred to as an *unpaired* image evaluation.

An appropriately designed separate image evaluation can decrease biases that might otherwise be introduced into an image assessment. As described in the first example below, an appropriately designed separate image evaluation should be performed when a goal of the study is to make comparative inferences about product performance (e.g., to compare the diagnostic performance of one medical imaging agent with another). As described in the second example, an appropriately designed separate image evaluation also can be used to demonstrate that a contrast agent contributes additional information to images obtained with the device alone.

Example 1: Comparative inferences of product performance

In a comparative study designed to show that the diagnostic performance of a new medical imaging agent is superior to that of an approved agent and that the new agent can substitute for the approved agent (see Section VI.D.1), an appropriate separate image evaluation of test images should be performed as the principal image analysis. The *test images* in this case are the images obtained with the new and the approved medical imaging agents. The two agents are not intended to be used together in actual clinical practice, and therefore the goal of such an *unpaired* image evaluation should be to show that the information obtained with the new agent is clinically and statistically superior to the information obtained with the approved agent. For any given patient, images obtained with the new agent should be evaluated independently of the evaluation of the images obtained with the approved agent, to the fullest degree practical.

If desired, a side-by-side (*paired)* comparison of images obtained with the new agent and the approved agent can be performed as a secondary image analysis. However, such a side-by-side comparison may yield estimates of diagnostic performance for the new agent that are biased. For example, in a side-by-side comparison of two medical imaging agents intended to detect masses, a blinded reader who sees an easily identifiable mass on an image obtained with the approved agent might be more likely to identify a mass on a juxtaposed image obtained with the new agent — even if that mass is not seen clearly on the latter image. In colloquial terms, the blinded reader may tend to *overread* the presence of masses on the image obtained with the new agent in such a paired comparison. Similarly, the blinded reader may tend to *underread* the image obtained with the new agent in a paired evaluation where a mass is not seen clearly on the image obtained with the approved agent.

Example 2: Contribution of additional information by a contrast agent

In a study intended to demonstrate that a contrast agent contributes additional information to images obtained with the device alone, it is often highly desirable to perform an appropriate separate image evaluation of test images as the principal image analysis (see the next section for an alternative approach). The *test images*, in this case, include both the images obtained before administration of contrast (the *unenhanced* images) and those obtained after administration of contrast (the *enhanced* images).

For example, in settings where the unenhanced image will not be used in clinical practice, the principal image analysis should be a separate image evaluation. The goal of such an unpaired image evaluation should be to show that the information obtained from the enhanced image is clinically and statistically superior to the information obtained from the unenhanced image. For any given patient, enhanced images obtained with the new agent should be evaluated independently of the evaluation of the unenhanced images, to the fullest degree practical.

b. Combined Image Evaluations

In a *combined* image evaluation, a reader simultaneously (or nearly simultaneously) evaluates two or more test images that were obtained under different conditions or at different times with respect to agent administration.[23] A combined image evaluation may resemble the conditions under which the product will be used clinically. For example, in some clinical situations both unenhanced and enhanced imaging studies are typically performed in patients. If so, such images often are evaluated concurrently in a comparative fashion.[24] However, as noted above, such combined image evaluations may increase the likelihood that bias will be introduced into the image evaluations (e.g., by systematic overreading or underreading particular findings on images).

A combined image evaluation can be performed by creating a set of combined images for each patient. These sets can then be presented to the blinded readers in random sequence. For example, in studies of contrast agents, both unenhanced and enhanced images can be obtained from each patient. The images for each patient, which were obtained at different times and under different conditions, may be viewed simultaneously by the blinded readers. Paired sets of images from different patients can be presented to the readers in random sequence.

When this type of reading is performed, however, an additional independent *separate* image evaluation should be completed on at least one of the members of the combination. Assuming that only one member is selected for this evaluation, the member chosen should be the member that usually is obtained under the current standard of practice (e.g., the unenhanced image). In this way, differences in the

[23] In the special case where only two test images are being evaluated, a *combined* image evaluation may also be referred to as a *paired* image evaluation.

[24] If images are evaluated only in a combined fashion, labeling of the medical imaging agent likely will specify that combined evaluations should be performed in clinical practice. If such labeling restrictions are not desired, then additional separate image evaluations should be performed.

evaluations of the combined reading with those of the separate reading can be assessed. When the goal is to show that the medical imaging agent adds information to images, these differences should demonstrate that the information from the combined images is clinically and statistically superior to information obtained from the separate image alone. The results of the combined and separate image evaluations can be analyzed statistically using paired comparisons.

For example, when a two-dimensional ultrasound study of blood vessels is performed with a microbubble contrast agent, a combined image evaluation could be performed by evaluating for each patient the unenhanced and enhanced images side-by-side (or in close temporal proximity). A separate independent evaluation of the unenhanced image of the blood vessel (i.e., images obtained with the device alone) for each patient could be performed. Assessing the differences for each patient between the results of the combined reading with those of the separate readings will allow the effects of the microbubble on the images to be determined.

As noted above, combined and separate image evaluations should be performed independently of one another (see Section VI.B.8.b). For example, to decrease recall bias, these combined evaluations should be designed to reduce the likelihood that the readers will be able to recall their assessment of the separate image assessment (or vice versa). Thus, different pages in the CRF should be used for the combined and separate evaluations, and the combined and separate image evaluations usually should be performed at different times without reference to prior results.

When differences between the combined and separate images are to be assessed, the combined CRF and separate CRF should contain items or questions that are identical so that differences can be calculated. On the separate CRF for a contrast agent, for example, the readers can be asked to rate on an ordinal scale (e.g., 0, 1, 2, 3, 4) the clarity of border delineation. The combined CRF should ask the same question and the difference in clarity should be calculated. The purpose of this approach is to reduce potential biases that may arise if the CRF contains only questions or items that ask for relative judgments. If desired, however, additional comparative questions and items can be added to the combined pages in the CRF. For example, the readers can be asked to rate the relative clarity of border delineation in the second image compared to the first (e.g., better, same, worse).

C. Truth Standards (Gold Standards)

A truth standard provides an independent way of evaluating the same variable being assessed by the investigational medical imaging agent. A truth standard is known or believed to give the true state of a patient or true value of a measurement. Truth standards are used to demonstrate that the results obtained with the medical imaging agent are valid and reliable. The following general principles should be incorporated prospectively into the design, conduct, and analysis of the major efficacy trials for medical imaging agents.

1. The true state of the subjects (e.g., diseased or nondiseased) should be determined with a truth standard without knowledge of the test results obtained with the medical imaging agent. In other words, the assessment of truth should be blinded to the imaging results with the medical imaging agent under study.

2. Conversely, test results obtained with the medical imaging agent should be evaluated without knowledge of the results obtained with the truth standard and without knowledge of outcome (see Section VI.B.7). In other words, evaluation of images obtained with the medical imaging agent should be blinded to the assessment of truth. Such evaluations decrease *diagnostic suspicion bias.*

3. Truth standards should not include as a component any test results obtained with the medical imaging agent (i.e., to avoid *incorporation bias)*. Similarly, the truth standard for contrast agents should not incorporate the results of the unenhanced image obtained with the device alone. In other words, the truth standard should be assessed independently of the imaging modality for which the medical imaging agent is intended because the features of the test image obtained with the medical imaging agent (e.g., the *enhanced image*) are likely to be correlated to the features of the image obtained with the device alone (e.g., the *unenhanced image*). For example, in the case of a CT contrast agent intended to visualize abdominal masses, unenhanced abdominal CT images generally should not be included in the truth standard. However, components of the truth standard might include results from other imaging modalities (e.g., MRI, ultrasonography).

4. Evaluation with the truth standard should be planned for all enrolled subjects, and the decision to evaluate a subject with the truth standard should not be affected by the test results with the medical imaging agent under study. For example, if patients with positive results with the test agent are evaluated preferentially with the truth standard (as compared to patients with negative test results), the results of the study may be affected by *partial verification bias*. Similarly, if patients with positive results with the test agent are evaluated preferentially with the truth standard and those with negative test results are evaluated preferentially with a less rigorous standard, the results of the study may be affected by *differential verification bias.*[25] Sponsors should seek FDA comment when it is anticipated that a meaningful proportion of enrolled

[25] Partial verification bias and differential verification bias are forms of *diagnostic work-up bias.*

subjects might not be evaluated with the truth standard or might be evaluated with a less rigorous standard. In such situations, it may be appropriate to evaluate clinical outcomes for the enrolled subjects (see Section IV.D.4).

From a practical perspective, diagnostic standards are derived from procedures that are considered more definitive in approximating the truth than the test agent. For example, histopathology or long-term clinical outcomes may be acceptable diagnostic standards for determining whether a mass is malignant. Diagnostic standards may not be error free, but for purposes of the clinical trial, they generally are regarded as definitive. It should be recognized, however, that misclassification of disease by the truth standard can lead to positive or negative biases in diagnostic performance measures (*misclassification bias*). The choice of the truth standard should be discussed with the Agency during design of the clinical trials to ensure that it is appropriate.

As noted in the rule for diagnostic radiopharmaceuticals, a valid assessment of actual clinical status can be provided by a diagnostic standard or standards of demonstrated validity. In the absence of such diagnostic standards, the actual clinical status can in some cases be established in another manner (e.g., through patient follow-up). However, when a suitable diagnostic standard is unavailable or cannot be assessed practically, consideration should be given to changing the focus of the study to evaluate the effects of the product on clinical outcomes (see Section IV.D.4).

Truth standards are typically other diagnostic tests (e.g., tissue biopsy to evaluate whether a mass is malignant), but truth standards also can be appropriate combinations of other clinical data and diagnostic tests. For example, a definitive determination about whether a patient enrolled in a clinical trial experienced an acute myocardial infarction could be obtained by evaluating the combination of patient history (e.g., nature and location of pain), 12-lead electrocardiogram (e.g., Q waves or not), and serum levels of cardiac enzymes (e.g., creatine phosphokinase) according to a prespecified algorithm. Using these data, a panel of experts that is blinded to the medical imaging results yielded by the test agent might then make the definitive determination about the presence or absence of disease (i.e., an acute myocardial infarction). In some cases, such as cases of suspected chronic infection or malignancy, the truth standard can involve obtaining clinical follow-up for a period following the imaging.

D. Comparison Groups

Clinical trials of medical imaging agents can include comparison groups for different purposes and can incorporate them into trial designs in a number of different ways. Before selecting comparison groups, discussions with the Agency are recommended.

1. Comparison to an Agent or Modality Approved for a Similar Indication

In the event that the test agent is being developed as an advance over an approved drug, biological product, or other diagnostic modality, a direct, concurrent comparison to the approved comparator(s) should be performed. The comparison should include an evaluation of both the safety and the efficacy data for the comparator(s) and the test agent. The image evaluation for the test product or modality should be done without knowledge of the imaging results obtained from the approved products or modalities (see Section VI.B.7).

Information from both test and comparator images should be compared not only to one another but also to an independent truth standard. This will facilitate an assessment of possible differences between the medical imaging agent and the comparator(s) and will enable comparative assessments of diagnostic performance. Such assessments could be obtained, for example, by comparing estimates of sensitivity, specificity, positive and negative predictive values, likelihood ratios, related measures, or receiver operating characteristic (ROC) curves for the different diagnostic agents. Note that two medical imaging agents could have similar values for sensitivity and specificity in the same set of patients, yet have poor agreement rates with each other. Similarly, two medical imaging agents could have good agreement rates, yet both have poor sensitivity and specificity values. In ROC analysis, overall areas under the curves obtained with different agents may be comparable, but areas under partial spans of the curves may be dissimilar. Likewise, one diagnostic agent may have superior diagnostic performance characteristics over another at one point on the ROC curve, but may have inferior diagnostic performance characteristics at a different point (see Section VII).

When a medical imaging drug or biologic is being developed for an indication for which other drugs, biological products, or diagnostic modalities have been approved, a direct, concurrent comparison to the approved drug, biologic, or diagnostic modality is encouraged. However, prior approval of a medical imaging agent for use in a particular indication does not necessarily mean that the results of a test with that agent can be used as a truth standard. For example, if a medical imaging agent has been approved on the basis of sufficient concordance of findings with truth as determined by histopathology, assessment of the new medical imaging agent should also usually include determination of truth by histopathology.

In studies that compare the effects of a test agent with another drug, biologic, or imaging modality, images taken before study enrollment with the comparator drug, biologic, or modality should not be used to determine whether a patient is enrolled in the study. These images also should not be part of the database used to determine test agent performance. Such baseline enrollment images have inherent selection bias because they are unblinded and based on referral and management preferences. All images used to determine the efficacy of the test agent and the comparator drug, biologic, or modality should be taken after study enrollment and within a time frame when the disease process is expected not to have changed significantly.

2. *Comparison to Placebo*

Whether the use of a placebo is appropriate in the evaluation of a medical imaging agent depends on the specific imaging agent, proposed indication, and imaging modality. In some cases, the use of placebos can help reduce potential bias in the conduct of the study and can facilitate unambiguous interpretation of efficacy or safety data. However, in some diagnostic studies (such as ultrasonography), products that are generally considered to be placebos (e.g., water, saline, or vehicle) can have some diagnostic effects. These should be used as controls to demonstrate that the medical imaging agent has an effect above and beyond that of the vehicle.

VII. STATISTICAL ANALYSIS

Statistical methods and the methods by which diagnostic performance will be assessed should be incorporated prospectively into the statistical analysis plan (see Section VI.B.2).

A. Statistical Methods

Many studies of imaging agents are designed to provide dichotomous, or ordered, categorical outcomes, and it is important that appropriate assumptions and statistical methods be applied in their analysis. Statistical tests for proportions and rates are commonly used for dichotomous outcomes, and methods based on ranks are often applied to ordinal data. Study outcomes can often be stratified in a natural way, such as by center or other subgroup category, and the Mantel-Haenszel[26] procedures provide effective ways to examine both binomial and ordinal data. Exact methods of analysis, based on conditional inference, should be employed when necessary. The use of model-based methods should also be encouraged. These models include logistic regression models for binomial data and proportional odds models for ordinal data. Log-linear models can be used to evaluate nominal outcome variables.

Dichotomous outcomes in studies that compare images obtained after the test agent to images obtained before the test agent are often analyzed as matched pairs, where differences in treatment effects can be assessed using methods for correlated binomial outcomes. These studies, however, may be problematic because they often do not employ blinding and randomization. For active- and placebo-control studies, including dose-response studies, crossover designs can often be used to gain efficiency. Subjects should be randomized to order of treatment. If subjects are not randomized to order of treatment, the order in which images

[26] For more on this topic, see Fleiss, Joseph, L., *Statistical Methods for Rates and Proportions*, 2nd ed., 1981, John Wiley and Sons, New York; and Woolson, Robert, *Statistical Methods for the Analysis of Biomedical Data*, 1987, John Wiley and Sons, New York.

are evaluated should be appropriately randomized. Study results from a crossover trial should always be analyzed with methods specifically designed for such trials.

B. Diagnostic Performance

Diagnostic validity can be assessed in a number of ways. For example, both with unenhanced and enhanced images, each could be compared to the truth standard, and the sensitivity and specificity of the unenhanced image could be compared to that of the enhanced image. Two different active agents can be compared in the same manner. Diagnostic comparisons can also be made when there are more than two outcomes to the diagnostic test results. Common methods used to test for differences in diagnosis include the McNemar test and the Stuart Maxwell test.[27] In addition, confidence intervals for sensitivity, specificity, and other measures should be provided in the analyses. Receiver operating characteristic analysis also may be useful in assessing the diagnostic performance of medical imaging agents over a range of threshold values.[28] For example, receiver operating characteristic analysis can be used to describe the relative diagnostic performance of two medical imaging agents if each test can be interpreted using several thresholds to define a positive (or negative) test result (see Section VI.D.1). For all planned statistical analyses, details of the analysis methods and specific hypotheses to be tested should be stated prospectively in the protocol as part of the statistical analysis plan. Sponsors should seek Agency comment on the design and statistical approach to analyses before the protocols are finalized.

VIII. GENERAL CONSIDERATIONS FOR SAFETY ASSESSMENTS OF MEDICAL IMAGING AGENTS

The safety evaluation of a medical imaging agent is generally similar to that of other drugs and biological products. However, in many cases, the special characteristics of medical imaging agents allow nonclinical and clinical safety assessments to be relatively efficient. The following sections discuss the special characteristics of a medical imaging agent that can lead to a more focused safety evaluation. These characteristics include its dose or mass, route of administration, frequency of use, and biological, physical, and effective half-lives.[29]

A. Dose or Mass

[27] Ibid.

[28] For an introduction to this topic, see Metz, Charles E. *Basic Principles of ROC Analysis,* Seminars in Nuclear Medicine 1978;VIII(4):283-298.

[29] See also the final rule on developing diagnostic radiopharmaceuticals. When a medical imaging agent does not possess any special characteristics, complete standard safety assessments should be performed.

Medical imaging agents can be administered at low mass doses. For example, the mass of a single dose of a diagnostic radiopharmaceutical often can be relatively small because device technologies can typically detect small amounts of a radionuclide. When a medical imaging agent is administered at a mass dose that is at the low end of the dose-response curve for adverse events, dose-related adverse events are less likely to occur.

B. Route of Administration

Some medical imaging agents are administered by routes that decrease the likelihood of systemic adverse events. For example, medical imaging agents that are administered as contrast media for radiographic examination of the gastrointestinal tract (e.g., barium sulfate) can be administered orally, through an oral tube, or rectally. In patients with normal gastrointestinal tracts, many of these products are not absorbed. Accordingly, systemic adverse events are less likely to occur in these patients. Therefore, after a sponsor demonstrates that such a product is not absorbed systemically in the population proposed for use, the product may be able to undergo a more efficient safety evaluation that primarily assesses local organ system toxicity, toxicities that are predictable (e.g., volume effects, aspiration), and effects after intraperitoneal exposure (e.g., after gastrointestinal perforation). However, if the product will be used in patients with gastrointestinal pathologies that increase absorption, additional nonclinical and clinical safety evaluations should be performed.

C. Frequency of Use

Many medical imaging agents, including both contrast agents and diagnostic radiopharmaceuticals, are administered relatively infrequently or as single doses. Accordingly, adverse events that are related to long-term use or to accumulation are less likely to occur with these agents than with agents that are administered chronically. Therefore, the nonclinical and clinical development programs for such products can generally omit long-term, or traditional, repeat-dose safety studies. That is, long-term repeat-dose toxicology studies (i.e., 3 months duration or longer) are normally not necessary for single-use agents.

However, in clinical settings where it is possible that the medical imaging agent will be administered repeatedly (e.g., to monitor disease progression), repeat-dose studies should be performed to assess safety and efficacy. Biological medical imaging agents are frequently immunogenic, and the development of antibodies after intermittent, repeated administration can alter the pharmacokinetics, biodistribution, safety, and/or imaging properties of such agents and, potentially, of immunologically related agents. Studies of immunogenicity in animal models are generally of limited value. Therefore, clinical data assessing the repeat use of a biological imaging agent should generally be obtained prior to application for licensure of such an agent.

D. Biological, Physical, and Effective Half-Lives

Diagnostic radiopharmaceuticals often use radionuclides with short physical half-lives or that are excreted rapidly. The biological, physical, and effective half-lives of diagnostic radiopharmaceuticals are incorporated into radiation dosimetry evaluations[30] that require an understanding of the kinetics of the distribution and excretion of the radionuclide and its mode of decay. Biological, physical, and effective half-lives should be considered in planning appropriate safety and dosimetry evaluations of diagnostic radiopharmaceuticals (see Sections IX and X.C).

IX. NONCLINICAL SAFETY ASSESSMENTS

The special characteristics of medical imaging agents described in Section VIII may allow for a more efficient nonclinical safety program. The nonclinical development strategy for an agent should be based on sound scientific principles; the agent's unique chemistry (including, for example, those of its components, metabolites, and impurities); and the agent's intended use. Sponsors are encouraged to consult with the Agency before submitting an IND application and during product development for recommendations and advice about the overall nonclinical development plan and proposed nonclinical protocols. In part, the number and types of nonclinical studies that should be conducted depend on the phase of the development, what is known about the agent or its pharmacologic class, its proposed use, and the indicated patient population.

In the discussion that follows, a distinction is made between biological products and drug products (see Section IX.A and Section IX.B, respectively). Existing specific guidance for biological products, which are typically evaluated in CBER, is referenced but not repeated here.

A. Nonclinical Safety Assessments for Biological Products

Many biological products raise relatively distinct nonclinical issues (e.g., immunogenicity and species specificity). To ensure consistency with Section 351 of the Public Health Service Act, the following Agency guidance documents should be reviewed on the preclinical evaluation of biological medical imaging agents:

- *S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals*, ICH, November 1997.

- *Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use*, February 1997.

---

[30] *Biological half-life* is the time needed for a human or animal to remove, by biological elimination, half of the amount of a substance that has been administered. *Effective half-life* is the time needed for a radionuclide in a human or animal to decrease its activity by half as a combined result of biological elimination and radioactive decay. *Physical half-life* is the time needed for half of the population of atoms of a particular radioactive substance to disintegrate to another nuclear form.

Sponsors are encouraged to consult with the appropriate CBER reviewing division for additional information when needed.

B. Nonclinical Safety Assessments for Drug Products (Non-Biological products)

The following sections describe ways in which nonclinical assessments of safety can be performed for contrast drug products and diagnostic radiopharmaceutical drug products. Sponsors are encouraged to consult with the appropriate CDER reviewing division for additional information when needed.

1. Contrast Agents

Because of the characteristics of contrast drug products and the way they are used, nonclinical safety evaluations of such drug products can be made more efficient with the following modifications:

- Long-term, repeat-dose toxicity studies in animals usually can be eliminated.
- Long-term rodent carcinogenicity studies usually can be omitted.[31]
- Reproductive toxicology studies often can be limited to an evaluation of embryonic and fetal toxicities in rats and rabbits and to evaluations of reproductive organs in other short-term toxicity studies.[32] However, a justification should be provided for any studies of reproductive toxicology that are not performed, and a formal request should be made to waive them.[33]

Additional safety considerations for contrast drug products can include the following: their large mass dose and volume (especially for iodinated contrast materials that are administered intravenously); osmolality effects; potential transmetalation of complexes of gadolinium, manganese, or iron (generally MRI drugs); potential effects of tissue or cellular accumulation on organ function (particularly if the drug is intended to image a diseased human organ system); and the chemical, physiological, and physical effects of ultrasound microbubble drugs (e.g., coalescence, aggregation, margination, and cavitation).

---

[31] Circumstances in which carcinogenicity testing may be recommended are summarized in the ICH guidance *S1A The Need for Long-Term Rodent Carcinogenicity Studies of Pharmaceuticals,* March 1996.

[32] See *S5A Detection of Toxicity to Reproduction for Medicinal Products* (ICH), September 1994, and *S5B Detection of Toxicity to Reproduction for Medicinal Products: Addendum on Toxicity to Male Fertility* (ICH), April 1996.

[33] Waiver regulations for INDs are set forth at 21 CFR 312.10; those for NDAs appear at 21 CFR 314.90.

2. *Diagnostic Radiopharmaceuticals*

Because of the characteristics of diagnostic radiopharmaceuticals and the way they are used, nonclinical safety evaluations of these drugs can be made more efficient by the following modifications:

- Long-term, repeat-dose toxicity studies in animals typically can be eliminated.

- Long-term rodent carcinogenicity studies usually can be omitted.

- Reproductive toxicology studies can generally be waived when adequate scientific justification is provided.[34]

- Because the radioactive component of the agent represents a likely genotoxic hazard, waivers for the performance of genotoxicity studies generally can be granted when adequate scientific justification is provided.[35]

In reproductive toxicology and genotoxicity studies, components other than the radionuclide should be considered separately because they may be genotoxins or teratogens, causing effects that may exceed those of the radioactivity alone.

Special safety considerations for diagnostic radiopharmaceuticals can include verification of the mass dose of the radiolabeled moiety; assessment of the mass, toxic potency, and receptor interactions for any unlabeled moiety; assessment of potential pharmacologic or physiologic effects due to molecules that bind with receptors or enzymes; and evaluation of all components in the final formulation for toxicity (e.g., excipients, reducing drugs, stabilizers, anti-oxidants, chelators, impurities, and residual solvents). An individual component should be tested if specific toxicological concerns are identified or if toxicological data for that component are lacking.

3. *Timing of Nonclinical Studies Submitted to an IND Application*

Appropriate timing of nonclinical studies should facilitate the timely conduct of clinical trials (including appropriate safety monitoring based on findings in nonclinical studies)

---

[34] See ICH *S5A* and ICH *S5B*.

[35] See *S2A Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals* (ICH), April 1996, and *S2B Genotoxicity: A Standard Battery for Genotoxicity Testing of Pharmaceuticals* (ICH), July 1997.

and should reduce the unnecessary use of animals and other resources.[36] The recommended timing of nonclinical studies for medical imaging drugs is summarized below.

a. Completed before phase 1

The following studies should be completed before phase 1:

- Safety pharmacology studies. Particular emphasis should be placed on human organ systems in which the medical imaging drug localizes and on organ systems that the product is intended to visualize, especially if the organ system has impaired function.

- Toxicokinetic and pharmacokinetic studies (see ICH guidances).

- Single-dose toxicity studies. *Expanded acute* single-dose toxicity studies are strongly recommended.[37] However, if short-term, repeated-dose toxicity studies have been completed, nonexpanded, single-dose toxicity studies may be sufficient.

  When repeated-dose toxicity studies have been performed but single-dose toxicology studies have not, dose selection for initial human studies will likely be based on the results of the no-adverse-effect level (NOAEL) obtained in the repeat-dose study. This will result in a dose selection for initial human administration that likely will be lower than otherwise would have been had dose selection been based on the results of acute, single-dose toxicity studies.

- For medical imaging drugs that are administered intravenously: (1) local tolerance and irritancy studies, including evaluations of misadministration or extravasation, (2) blood compatibility studies, including evaluations of hemolytic effects, and (3) effects on protein flocculation.

- Radiation dosimetry, if applicable.

- In vitro genotoxicity studies (see Section IX.B.2 for diagnostic radiopharmaceuticals).

[36] See *M3 Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals* (ICH), July 1997.

[37] See *Single Dose Acute Toxicity Testing for Pharmaceuticals*, August 1996.

b. Completed before phase 2

The following studies should be completed before phase 2:

- Short-term, repeated-dose toxicity studies.
- Immunotoxicity studies.
- In vivo genotoxicity studies (see Section IX.B.2 for diagnostic radiopharmaceuticals).

c. Completed before phase 3

Reproductive toxicity studies should be completed before phase 3, if needed (see Section IX.B.2 for diagnostic radiopharmaceuticals).

d. Completed no later than the end of phase 3

The following studies should be completed no later than the end of phase 3:

- Drug interaction studies.
- In vivo or in vitro studies that further investigate adverse effects seen in previous nonclinical studies.

X. CLINICAL SAFETY ASSESSMENTS[38]

Indications for medical imaging drugs or biological products should be supported with information demonstrating that the potential benefits of the use of the medical imaging agent outweigh the potential risks to the patient. Potential risks include both the risks related to administration of the agent and the risks of incorrect diagnostic information. Incorrect diagnostic information includes, but is not limited to, inaccurate structural, functional, physiological, or biochemical information; false positive or false negative diagnostic determinations; and information leading to inappropriate decisions in diagnostic or therapeutic management (see Section IV.A). FDA weighs these potential benefits and potential risks when making its decision about whether to approve a marketing application (e.g., NDA or BLA).

The special characteristics of medical imaging agents described in section VIII may allow for a more efficient clinical safety program. Moreover, this guidance defines two categories for medical imaging

[38] The final rule, "Expedited Safety Reporting Requirements for Human Drug and Biological Products," October, 7, 1997 (62 FR 52237).

agents: Group 1 and Group 2. The extent of clinical safety monitoring and evaluation differs for these two categories. Medical imaging agents classified by FDA as *Group 1 medical imaging agents* will usually be able to undergo a more focused clinical safety evaluation during development. Those agents classified by the Agency as *Group 2 medical imaging agents* should undergo standard clinical safety evaluations in clinical trials throughout development.

In the discussion that follows, *standard clinical safety evaluations* include serial assessments of patient symptoms, physical signs, clinical laboratory tests (e.g., blood chemistry, hematology, coagulation profiles, urinalyses), other tests (e.g., electrocardiograms as appropriate), and adverse events. Additional specialized evaluations should be performed when appropriate (e.g., immunological evaluations, creatine kinase isoenzymes) or if a particular toxicity is deemed possible based on animal studies or the known chemical or pharmacological properties of the medical imaging agent. These standard clinical safety evaluations can be tailored based on the characteristics of the medical imaging agent under study (e.g., dose, route of administration, frequency of use, and biological half-life), and on the results of nonclinical safety assessments and the results of clinical pharmacokinetic/biopharmaceutics studies (see Sections VIII and IX). The duration of clinical monitoring should be sufficient to identify possible effects that may lag behind those predicted by pharmacokinetic analyses.

If some of these standard clinical safety evaluations are felt to be unnecessary, this should be discussed with the reviewing division. Sponsors should seek FDA comment on the clinical safety monitoring plans in clinical studies before such studies are initiated.

Note that under the safety-margin criteria described below (see Section X.A.1.a), medical imaging agents that are administered in low mass doses to humans (e.g., diagnostic radiopharmaceuticals) usually are more likely to be given a Group 1 designation than those administered in higher mass doses.[39] There are important exceptions, including cases where the medical imaging agents are likely to be immunogenic (e.g., biological products) or when medical imaging agents cause adverse reactions that are not dose-related (e.g., idiosyncratic drug reactions).

A. Group 1 Medical Imaging Agents

A medical imaging agent can be classified as a Group 1 medical imaging agent if it meets the following three conditions:

- The medical imaging agent meets *either* the safety-margin criteria or the clinical-use criteria described below (see Sections X.A.1 and X.A.2, respectively).

[39] Groups 1 and 2 include diagnostic radiopharmaceuticals. This classification conforms with the final rule for diagnostic radiopharmaceuticals, which states that diagnostic radiopharmaceuticals may be categorized based on defined characteristics related to their risk.

! The medical imaging agent is not a biological product.[40, 41]

! The medical imaging agent is not a diagnostic radiopharmaceutical containing a radionuclide that emits alpha or beta particles.[42, 43]

Standard clinical safety evaluations should be performed in all clinical investigations of medical imaging agents until a Group 1 designation is assigned. Once a medical imaging agent is granted a Group 1 designation, reduced human safety monitoring in subsequent human trials can be planned. For example, human safety monitoring may be limited to recording adverse events and monitoring particular organs or tissues of interest for toxicity (such as organs that showed toxicity in the animal studies or the tissues in which the medical imaging agent localizes).

A Group 1 designation can be retained throughout product development if safety concerns are not raised subsequently in nonclinical and clinical studies. If safety concerns are identified, the medical imaging drug can be given a Group 2 designation for the remainder of product development.

1. Safety-Margin Criteria

Under the safety-margin criteria, medical imaging agents can obtain a Group 1 designation if, as described below, the results of nonclinical studies *and* initial human experience both are consistent with the conditions summarized in the following two subsections (i.e., Sections X.A.1.a and X.A.1.b):

a. Results of nonclinical studies

To obtain a Group 1 designation under the safety-margin criteria, a medical imaging agent should have an adequately documented margin of safety as assessed in the nonclinical studies outlined in the following list:[44]

---

[40] Medical imaging products that are biological products, such as radiolabeled cells, monoclonal antibodies, or monoclonal antibody fragments, will not normally be classified as Group 1 medical imaging agents because of their potential to elicit immunologic responses.

[41] See also the final rule, "Adverse Experience Reporting Requirements for Licensed Biological Products," (59 FR 54042; October 27, 1994).

[42] This statement does not apply to diagnostic radiopharmaceuticals that are pure positron ($\$^+$) emmiters.

[43] Group 1 diagnostic radiopharmaceuticals may include radionuclides, ligands, and carriers that are known to be biologically inactive. This group may include radionuclides, ligands, and carriers used at radiation doses or mass dosages that are similar to, or less than, those used previously. This group also may include radionuclides, ligands, and carriers that have been documented not to produce adverse reactions.

! The no-observed-adverse-effect level (NOAEL)[45] in expanded-acute, single-dose toxicity studies in suitable animal species should be at least one hundred times (100x) greater than the maximal dose and dosage to be used in human studies. Such expanded, acute, single-dose toxicity studies should be completed before the medical imaging agent is introduced into humans (see Section IX.B.3).

! The NOAEL in safety pharmacology studies in suitable animal species should be at least one hundred times (100x) greater than the maximal dose and dosage to be used in human studies. Such safety pharmacology studies should be completed before the medical imaging agent is introduced into humans (see Section IX.B.3).

! The NOAEL in short-term, repeated-dose toxicity studies in suitable animal species should be at least twenty-five times (25x) greater than the maximal dose and dosage to be used in human studies.[46] Such short-term, repeated-dose toxicity studies can be performed either before the medical imaging agent is introduced into humans, or concurrently with early human studies, but should be completed before phase 2 (see Section IX.B.3).

To establish these margins of safety, the NOAELs should be assessed in properly designed and conducted studies and should be appropriately adjusted. *Appropriately adjusted* means that dosage comparisons between animals and humans should be suitably modified for factors such as body size (e.g., body surface area) and otherwise adjusted for possible pharmacokinetic and toxicokinetic differences between animals and humans (e.g., differences in absorption for products that are administered orally).

---

[44] In addition, the medical imaging agent should meet the conditions described for the results of initial human experience (see Section X.A.1.b):

[45] For purposes of classification into Groups 1 and 2 in this section of this Guidance, the no-observed-adverse-effect-level (NOAEL) is defined as the highest dose tested in animals with no adverse effects. In this context, an adverse effect is an event that is reasonably serious and would be unacceptable if produced by the initial dose of a test agent in a phase 1 clinical trials conducted in healthy volunteers.

[46] Short-term, repeated-dose toxicity studies may identify toxicities associated with accumulation of a medical imaging agent or its metabolites. In addition, even if such accumulation is not anticipated (e.g., non-metabolized medical imaging agents with short half-lives), short-term repeated-dose toxicity studies may identify toxicities caused by repeated toxic insults, each of which may be below the threshold of detection in expanded-acute, single-dose toxicity studies.

Note that medical imaging agents granted a Group 1 designation should undergo other nonclinical toxicological studies as described in Section IX, such as genotoxicity, reproductive toxicity, irritancy studies, and drug-drug interaction studies.

i. Possible exceptions to the safety margins

A Group 1 designation may be possible for some medical imaging agents when the NOAELs are slightly less than the multiples specified above. Such designations will be determined on a case-by-case basis. These determinations will take into consideration, among other things, how close the NOAELs are to the multiples specified above, the amount of safety information known about chemically similar and pharmacologically related medical imaging agents, the nature of observed animal toxicities, and whether adverse events have occurred during initial human experience, including the nature of such adverse events (see Section X.A.1.b).

ii. Formulations Used in Nonclinical Studies

The formulation used to establish safety margins in these nonclinical studies should be identical, to the fullest extent practical, to the formulation to be used in clinical trials and that is intended for marketing. Any differences in the formulations used in the clinical trials and nonclinical studies should be specified so that any impact on the adequacy of the nonclinical studies can be determined. In some cases, it may be infeasible or impractical to administer the intended clinical formulation to animals in multiples of the maximal human dose that were specified above (e.g., the volume of such an animal dose may be excessive). In these cases, alternative strategies can be employed, such as dividing the daily dose (e.g., into a morning and evening dose), or by using a more concentrated formulation of the medical imaging agent. In cases when such alternative strategies are infeasible or impractical, the maximal feasible daily dose can be administered. If alternative dosing strategies or use of a maximal feasible daily dose are being contemplated, sponsors are encouraged to discuss their plans with FDA before studies are initiated.

b. Results of initial human experience

To obtain a Group 1 designation under the safety-margin criteria, the following conditions should be met (in addition to the conditions described above for the results of nonclinical studies):

! Safety issues should not be identified during initial human use of the medical imaging agent in appropriately designed studies that include adequate and documented standard clinical safety evaluations. That is, given the multiples that were specified in animal studies for a Group 1 designation (see Section X.A.1.a), identification of any adverse event during initial human use would be considered significant, particularly if those adverse events were not predicted from effects observed in animals. If adverse events occur at any time during human studies, the medical imaging agent may be reclassified as a Group 2 medical imaging agent.

! Human pharmacokinetic studies should be performed in phase 1 to allow adequate comparisons of exposure to be made between humans and the species used in the nonclinical studies. Such pharmacokinetic data can allow a more meaningful assessment of the relevance of the animal safety data (e.g., toxicokinetics).

FDA anticipates that most Group 1 designations based on the safety-margin criteria will occur at the end of phase 1, after animal studies and initial human trials have been completed and after all the conditions specified under this section have been met.

2. Clinical Use Criteria

Another way that medical imaging agents can obtain a Group 1 designation is by adequately documenting extensive prior clinical use during which safety issues were not identified. This means that human toxicity or adverse events should not have been observed during prior human use of the medical imaging agent when clinical doses (including both mass and radiation doses, if applicable) of the agent were administered under conditions where adequate safety monitoring was performed and the lack of human toxicity was adequately documented. For example, previous human use of such a medical imaging agent at relevant doses should not have been associated with adverse events and should not have been associated with effects with potential clinical consequences.[47] The methods used to monitor for adverse events and effects with potential clinical consequences should be documented to ensure that monitoring would have been able to detect such adverse events and responses had they been present.

Group 1 designations based on the clinical-use criteria can occur at any time during drug development (e.g., after the conditions specified in this section have all been met).

---

[47] In this context, effects with potential clinical consequences include pharmacologic, physiologic, biochemical or structural activities that need not necessarily be adverse or toxic. However, localization of a medical imaging agent in a target organ or target tissue (e.g., by binding to a tissue receptor) is not considered by itself to be a pharmacologic, physiologic, biochemical, structural, or toxic effect, unless such localization produces perturbations that are clinically demonstrable. Similarly, because these agents are intended for use in medical imaging, the ability to detect a medical imaging agent in a target organ or tissue by the intended imaging modality is not considered by itself to be a pharmacologic, physiologic, biochemical, structural, or toxic effect.

B. Group 2 Medical Imaging Agents

Group 2 medical imaging agents are medical imaging drugs or biological products that do not meet the criteria for Group 1 medical imaging agents. Group 2 medical imaging agents have been shown to be, or can be presumed to be (e.g., biological products) biologically active in animal studies or in human studies when administered at dosages that are similar to those intended for clinical use. Group 2 diagnostic radiopharmaceuticals are a subset of this group.[48] For Group 2 medical imaging agents, standard clinical safety evaluations and monitoring should be performed in clinical trials.

C. Radiation Safety Assessment for All Diagnostic Radiopharmaceuticals[49]

Radiation safety assessments should be fully documented for both Group 1 and Group 2 diagnostic radiopharmaceuticals. The radiation safety assessment should establish the radiation dose of a diagnostic radiopharmaceutical by radiation dosimetry evaluations in humans and appropriate animal models. Such an evaluation should consider dosimetry to the total body, to specific organs or tissues (including critical or sensitive organs or tissues), and, as appropriate, to target organs or target tissues. The radiation doses of diagnostic radiopharmaceuticals should be kept as low as reasonably achievable (ALARA). The maximum tolerated radiation dose need not be established. For diagnostic radiopharmaceuticals, estimates of the organ dosimetry should be performed in animals prior to the first phase 1 study. Phase 1 studies of diagnostic radiopharmaceuticals should include studies that will obtain sufficient data for dosimetry calculations (21 CFR 312.23(a)(10)(ii)).

1. General Considerations

An IND sponsor should submit sufficient data from animal or human studies to allow a reasonable calculation of radiation absorbed dose to the whole body and to critical organs upon administration to a human subject (21 CFR 312.23(a)(10)(ii)). At a minimum, the following organs and tissues should be included in dosimetry estimates: (1) all target organs/tissues; (2) bone; (3) bone marrow; (4) liver; (5) spleen; (6) adrenal glands; (7) kidney; (8) lung; (9) heart; (10) urinary bladder; (11) gall bladder; (12) thyroid; (13) brain; (14) gonads; (15) gastrointestinal tract; and (16) adjacent organs of interest. When a diagnostic radiopharmaceutical is being developed for pediatric use, it may be appropriate to evaluate the radiation absorbed dose in all

---

[48] Group 2 diagnostic radiopharmaceuticals can also include radionuclides and carriers that are known to be biologically active. This group includes radionuclides and carriers used at radiation doses or mass dosages that are higher than those used previously, including radionuclides and carriers that have been documented to produce adverse reactions.

[49] This section is based largely on the radiation dosimetry section of CBER's *Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use,* February 1997.

organs, rather than in selected organs. Moreover, dosimetry evaluations should be performed in the pediatric age groups (e.g., neonates, infants, children, adolescents) in which the diagnostic radiopharmaceutical is intended to be used.

The amount of radiation delivered by internal administration of diagnostic radiopharmaceuticals should be calculated by internal radiation dosimetry. The absorbed fraction method of radiation dosimetry has been described by the Medical Internal Radiation Dose (MIRD) Committee of the Society of Nuclear Medicine and the International Commission on Radiological Protection (ICRP).

The methodology used to assess radiation safety should be specified including reference to the body models that were used. The mathematical equations used to derive the radiation doses and the absorbed dose estimates should be provided along with a full description of assumptions that were made. Sample calculations and all pertinent assumptions should be listed and submitted. The reference to the body, organ, or tissue model used in the dosimetry calculations should be specified, particularly for new models being tested.

Safety hazards for patients and health care workers during and after administration of the radiolabeled antibody should be identified, evaluated, and managed appropriately.

2. Calculation of Radiation Dose to the Target Organs or Tissues

The following items should be determined based on the average patient:

a. The amount of radioactivity that accumulates in the target tissue(s) or organ(s)

b. The amount of radioactivity that accumulates in tissues adjacent to the target tissue(s) or organ(s)

c. The residence time of the diagnostic radiopharmaceutical in the target tissue(s) or organ(s) and in adjacent regions

d. The radiation dose from the radionuclide, including the free radionuclide and any daughter products generated by decay of the radionuclide

e. The total radiation dose from bound, free, and daughter radionuclides associated with the diagnostic radiopharmaceutical, based on immediate administration following preparation and upon delayed administration at the end of the allowed shelf life

3. Maximum Absorbed Radiation Dose

The amount of radioactive material administered to human subjects should be the smallest radiation dose that is practical to perform the procedure without jeopardizing the benefits obtained.

a. The amount of radiation delivered by the internal administration of diagnostic radiopharmaceuticals should be calculated by internal radiation dosimetry using either the MIRD or ICRP methods.

b. Because of known or expected toxicities associated with radiation exposure, dosimetry estimates should be obtained as described above.

c. Calculations should anticipate possible changes in dosimetry that might occur in the presence of diseases in organs that are critical in metabolism or excretion of the diagnostic radiopharmaceutical. For example, renal dysfunction may cause a larger fraction of the administered dose to be cleared by the hepatobiliary system (or vice versa).

d. Possible changes in dosimetry resulting from patient-to-patient variations in antigen or receptor mass should be considered in dosimetry calculations. For example, a large tumor mass may result in a larger than expected radiation dose to a target organ from a diagnostic radiopharmaceutical that has specificity for a tumor antigen.

e. The mathematical equations used to derive the estimates of the radiation dose and the absorbed dose should be provided along with a full description of assumptions that were made. Sample calculations and all pertinent assumptions should be listed.

f. Calculations of dose estimates should be performed assuming freshly labeled material (to account for the maximum amount of radioactivity) as well as the maximum shelf life of the diagnostic radiopharmaceutical (to allow for the upper limit of radioactive decay contaminants). These calculations should (1) include the highest amount of radioactivity to be administered; (2) include the radiation exposure contributed by other diagnostic procedures such as roentgenograms or nuclear medicine scans that are part of the study; (3) be expressed as gray (Gy) per megabecquerel (MBq) or per millicurie (mCi) of radionuclide; and (4) be presented in a tabular format and include doses of individual absorbed radiation for the target tissues or organs and the organs listed above in Section X.C.1.

GLOSSARY

Note: Subjects in trials of medical imaging agents are often classified into one of four groups depending on (1) whether disease is present (often determined with a truth standard or *gold standard*) and (2) the results of the diagnostic test of interest (positive or negative). The following table identifies the variables that are used to estimate the parameters defined below.

Test Result:	**Disease:** Present (+)	Absent (-)	
Positive (+)	**a** true positive=TP	**b** false positive=FP	**m1 = a+b** = TP+FP total with positive test
Negative (-)	**c** false negative=FN	**d** true negative=TN	**m2 = c+d** = FN+TN total with negative test
	n1 = a+c = TP+FN total with disease	**n2 = b+d** = FP+TN total without disease	**N = a+b+c+d** = TP+FP+FN+TN total in study

Accuracy: (1) In common usage, *accuracy* is the quality of being true or correct. (2) As a measure of diagnostic performance, *accuracy* is a measure of how faithfully the information obtained using a medical imaging agent reflects reality or *truth* as measured by a truth standard or *gold standard*. Accuracy is the proportion of cases, considering both positive and negative test results, for which the test results are correct (i.e., concordant with the truth standard or *gold standard*). Accuracy = (a+d)/N = (TP+TN)/(TP+FP+FN+TN).

Likelihood ratio: A measure that can be interpreted either as (a) the relative *odds* of a diagnosis, such as being diseased or nondiseased, for a given test result, or (b) the relative *probabilities* of a given test result in subjects with and without the disease. This latter interpretation is analogous to a relative risk or risk ratio.

1. For tests with dichotomous results (e.g., positive or negative test results), the likelihood ratio of a positive test result can be expressed as LR(+), and the likelihood of a negative test result can be expressed as LR(-). See the equations below:

$$LR(+) = \frac{\frac{a}{n1}}{\frac{b}{n2}} = \frac{sensitivity}{1 - specificity} = \frac{TruePositiveRate}{FalsePositiveRate} = \frac{\frac{a}{b}}{\frac{n1}{n2}} = \frac{PostTestOdds(+)}{PreTestOdds}$$

$$LR(-) = \frac{\frac{c}{n1}}{\frac{d}{n2}} = \frac{1 - sensitivity}{specificity} = \frac{FalseNegativeRate}{TrueNegativeRate} = \frac{\frac{c}{d}}{\frac{n1}{n2}} = \frac{PostTestOdds(-)}{PreTestOdds}$$

LR(+): *Interpreted as relative odds:* LR(+) is the post-test odds of the disease (among those with a positive test result) compared to the pretest odds of the disease.

Interpreted as relative probabilities: LR(+) is the probability of a positive test result in subjects with the disease compared to the probability of a positive test result in subjects without the disease.

LR(-): *Interpreted as relative odds*: LR(-) is the post-test odds of the disease (among those with a negative test result) compared to the pretest odds of the disease.

Interpreted as relative probabilities: LR(-) is the probability of a negative test result in subjects with the disease compared to the probability of a negative test result in subjects without the disease.

2. For tests with several levels of results, such as tests with results expressed on ordinal or continuous scales, the likelihood ratio can be used to compare the proportions of subjects with and without the disease at different levels of the test result. Alternatively, the likelihood ratio can be used to compare the post-test odds of disease at a particular level of test result compared with the pretest odds of disease. Thus, the generalized likelihood ratio can reflect diagnostic information at any level of the test result.

Negative predictive value: The probability that a subject does not have the disease given that the test result is negative. Synonyms include *predictive value negative.* Negative predictive value = d/m2 = TN/(TN+FN).

By application of Bayes' Rule, the negative predictive value also can be defined as a function of pretest probability of disease (p), sensitivity, and specificity:

Negative predictive value = [(1-p) C specificity]/[(1-p) C specificity + p C (1-sensitivity)]

Odds: The probability that an event will occur compared to the probability that the event will not occur. Odds = (probability of the event)/(1 - probability of the event).

Positive predictive value: The probability that a subject has disease given that the test result is positive. Synonyms include *predictive value positive.* Positive predictive value = a/m1 = TP/(TP+FP)

By application of Bayes' Rule, the positive predictive value also can be defined as a function of pretest probability of disease (p), sensitivity, and specificity:

Positive predictive value = (p C sensitivity)/[p C sensitivity + (1-p) C (1-specificity)]

Post-test odds of disease: The odds of disease in a subject after the diagnostic test results are known. Synonyms include *posterior odds of disease*. For subjects with a positive test result, the post-test odds of disease = a/b = TP/FP. For subjects with a negative test result, the post-test odds of disease = c/d = FN/TN. The following expression shows the general relationship between the post-test odds and the likelihood ratio: Post-test odds of disease = Pretest odds of disease x Likelihood ratio.

Post-test probability of disease: The probability of disease in a subject after the diagnostic test results are known. Synonyms include *posterior probability of disease*. For subjects with a positive test result, the post-test probability of disease = a/m1 = TP/(TP+FP). For subjects with a negative test result, the post-test probability of disease = c/m2 = FN/(TN+FN).

Precision: A measure of the reproducibility of a test, including reproducibility within and across doses, rates of administration, routes of administration, timings of imaging after product administration, instruments, instrument operators, patients, and image interpreters, and possibly other variables. Precision is usually expressed in terms of variability, using such measures as confidence intervals and/or standard deviations. Precise tests have relatively narrow confidence intervals (or relatively small standard deviations).

Pretest odds of disease: The odds of disease in a subject before doing a diagnostic test. Synonyms include *prior odds of disease*. Pretest odds of disease = n1/n2 = (TP+FN)/(TN+FP).

Pretest probability of disease: The probability of disease in a subject before doing a diagnostic test. Synonyms include *prevalence of disease* and *prior probability of disease*. Pretest probability of disease = n1/N = (TP+FN)/(TP+FP+FN+TN).

Probability: The likelihood of occurrence of an event, expressed as a number between 0 and 1 (inclusive).

Receiver operating characteristic (ROC) curve: A graphical representation of pairs of values for *true positive rate* (or sensitivity) and the corresponding *false positive rate* (or 1-specificity) for a diagnostic test. Each pair is established by classifying the test result as *positive* when the test outcome equals or exceeds the value set by a given threshold, and *negative* when the test outcome is less than this threshold value. For example, if a five-point ordinal scale is used to rate the likelihood of malignancy for a tumor (e.g., definitely benign, probably benign, equivocal, probably malignant, definitely malignant), setting the threshold at *equivocal* will classify tumors as malignant (i.e., a *positive* test result) when the test outcome is at this level or higher and will classify tumors as nonmalignant (i.e., a *negative* test result) when the test outcome is less than this level. To generate an ROC curve, the sensitivity and specificity of the diagnostic test are calculated and graphed for several thresholds (e.g., all values of the rating scale). In a typical ROC curve, values for *true positive rate* (or sensitivity) are plotted on the vertical axis, and the corresponding values for *false positive rate* (or 1-specificity) are plotted on the horizontal axis.

Sensitivity: The probability that a test result is positive given the subject has the disease. Synonyms include *true positive rate.* Sensitivity = a/n1 = TP/(TP+FN).

Specificity: The probability that a test result is negative given that the subject does not have the disease. Synonyms include *true negative rate.* Specificity = d/n2 = TN/(TN+FP).

Truth standard (gold standard): An independent method of measuring the same variable being measured by the investigational drug or biologic that is known or believed to give the *true* value of the measurement.